

Social Research Methods Review Vol. 2

社会研究方法评论

主编 赵联飞 赵锋

第2卷



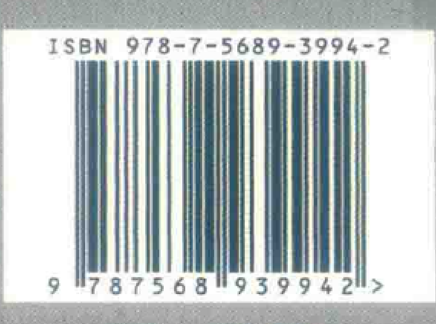
重庆大学出版社

社会研究方法评论

Social Research Methods Review Vol. 2







定价：42.00元

社会研究方法评论

Social Research Methods Review Vol. 2

第2卷

主编　赵联飞　赵锋

重庆大学出版社

图书在版编目(CIP)数据

社会研究方法评论. 第2卷 / 赵联飞, 赵锋主编. --
重庆: 重庆大学出版社, 2023.7
(万卷方法)

ISBN 978-7-5689-3994-2

Ⅰ.①社… Ⅱ.①赵… ②赵… Ⅲ.①社会学—研究
方法 Ⅳ.①C91-03

中国国家版本馆 CIP 数据核字(2023)第 109455 号

社会研究方法评论 第2卷

SHEHUI YANJIU FANGFA PINGLUN DIERJUAN

主 编 赵联飞 赵 锋
策划编辑:林佳木
责任编辑:石 可 版式设计:林佳木
责任校对:王 倩 责任印制:张 策

*

重庆大学出版社出版发行
出版人:饶帮华
社址:重庆市沙坪坝区大学城西路 21 号
邮编:401331
电话:(023)88617190 88617185(中小学)
传真:(023)88617186 88617166
网址:http://www.cqup.com.cn
邮箱:fxk@cqup.com.cn(营销中心)
全国新华书店经销
重庆华林天美印务有限公司印刷

*

开本:890mm×1240mm 1/32 印张:7.5 字数:182千
2023年7月第1版 2023年7月第1次印刷
ISBN 978-7-5689-3994-2 定价:42.00元

栏目

2022年2月创刊　　总第2卷

特稿

论文

方法评论

田野与课堂

铸器坊

新知与书评

目录

沈崇麟．五城市调查最终调查数据产生始末究[M/OL]//赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

五城市调查最终调查数据产生始末

沈崇麟[1]

摘要：本文回顾了1981—1983年开展的五城市调查的起源和组织过程，对五城市调查的社会影响进行了回顾，并介绍了五城市调查数据再次整理的过程。

关键词：五城市调查；数据

Abstract: This paper reviews the origin and organizational process of the Five Cities Survey conducted from 1981 to 1983, reviews the social impact fund of the Five Cities Survey, and introduces the process of re-collating the data of the Five Cities Survey.

Key words: Five Cities Survey

---

[1] 作者简介：沈崇麟，中国社会科学院社会学研究所研究员、社会学系教授（退休），曾任社会调查与方法研究室主任（1992—2004），社会调查与数据处理研究中心主任（1992—2017）。

一、五城市调查的起源

中国社会学重建之后于1981—1983年进行的五城市调查，是中国社会学抽样调查的起点。它始于1981年在天津尖山街进行的婚姻家庭和生育问题调查。那时中国社会学会接待了受美中学术交流学会资助来华访问的美国人类学家，美国纽约市立大学人类学教授帕斯特内克（Burton Pasternak）。应他的要求，费孝通先生安排他到天津进行实地调查。调查的题目是婚姻家庭和生育问题，调查的对象为调查所在地全部55岁以上的已婚妇女。调查采取的是入户访谈式的问卷调查。中国社会科学院社会学研究所的潘乃谷和天津社会科学院社会学研究所的潘允康组织实施了此项调查。

当时，费孝通先生认为，中国社会学正处在起步阶段，百废待兴，要认真学习西方国家的经验。他认为西方的经验不仅限于理论方法，同时也包括如何做调查。他在天津调查结束后对那时所里正在北京宣武区椿树街道的东河沿居委会进行社会调查基地建设的人（我和所里的同事薛寅、张仙桥和王荣芬等）说："不入虎穴焉得虎子。"他提出要在东河沿做一个同样的调查，把他们的经验学到手。于是有了1982年的"东河沿婚姻家庭和生育调查"。在接受任务以后我们参照天津调查的问卷设计了东河沿调查的问卷。为了有可比性，我们保留了天津调查绝大部分项目，但是也根据研究的需要，增加了不少新的项目，所以东河沿调查的内容比尖山街问卷调查篇幅长不少。调查对象也从55岁以上的已婚妇女，改为全体已婚妇女。和天津调查一样，东河沿调查也采用入户访谈的方式。

整个东河沿调查都是在费孝通先生的指导下进行的。参加调查

的人几次到他家中开会，那时他的家还在中央民族学院内，是一个两居室的单元房。地方狭窄，只有两个卧室。他就坐在卧室中的一张八仙桌旁，听取我们的汇报并对调查工作作出指示。他提出调查人员不仅要调查，而且要利用调查得到的资料写文章。根据他的要求，我用东河沿的资料，写了一篇论婚姻的门当户对的论文，他亲自审阅并批示，认为论文很好，有理有据，指出社会学者论文就是应该从调查资料出发来写文章。在此期间，费孝通先生还在家中安排了对《中国青年生育意愿调查》的讨论。希望我们能利用东河沿调查的资料对它做一些补充。实际上因为那时《中国青年生育意愿调查》这一项目的实地调查已经完成，进入了分析写作的阶段，所以东河沿调查的资料实际上并没有起到什么补充作用，与会的人只是对分析和写作的框架提了一些意见，这些意见中张之毅先生的意见是非常有分量的，对最终报告的完成具有指导意义。

如果说东河沿调查对家庭研究有什么贡献的话，那就是在东河沿调查中，我们第一次提出了用与被调查人共同生活这一指标作为“家庭”这一概念的操作定义。在东河沿调查最初进行的访谈中，我们发现在问及家庭人口的时候，如果不加任何限制，被调查人常常会将调查时健在的夫方所有的男性亲属及其配偶包括进来，家庭人口一般都会有十几口，有时甚至会有几十口。这固然反映了当时中国人心目中“家”这一概念，但显然这不是学术上的“家庭”的概念。因此，在遇到这样的情况时我都会进一步询问：“在这些人中，哪些人现在和你一起过日子？”这时被调查人的家庭人口便会迅速减少，因为那些不在一起生活的人此时都会被排除出去，这就使家庭人口回归到“家”的理论定义限定的范围。在调查中我们还发现，对“共同生活”这一概念我们不必作出硬性规定，如是否住

在一起、是否在一起吃饭等，只要由被调查人主观认定就可以了。换言之，东河沿调查既没有使用家庭社会学经典的关于“屋顶”（one roof one family）的操作定义，也没有使用费孝通先生“一起吃饭”（同灶）这样的标准，而是采用了主观认定的“共同生活”这一标准作为家庭的操作定义。之后其他调查点进行的调查，以及我之后所做的家庭调查都采用它作为家庭的操作定义。东河沿调查使我认识到在家庭研究中怎样在操作层面定义一个“家”，是一个十分重要的问题。

二、联合调查组的成立

1982年夏，中国社会学会第一届年会在武汉的华中工学院，即今天的华中科技大学举行。我应费孝通先生之命，在会上做了一个题为《论基地》的报告。关于“调查基地”的思想是费孝通先生一贯提倡和坚持的社会学者做学问之道。他认为社会学者不能脱离社会，必须时刻关注社会的变化，而社会调查基地正是社会学者观察和了解社会的窗口。正是在费孝通先生的倡导和主持下，社会学恢复之后的第一个调查基地得以建立——北京宣武区椿树街道东河沿居委会调查基地，即东河沿调查基地。而东河沿婚姻家庭和生育调查则是在那个基地完成的第一个调查项目。所以我在年会上的报告除了阐述自己对费孝通先生关于社会调查基地思想的理解以及在东河沿进行社会调查基地建设的情况外，也以相当的篇幅对东河沿的婚姻家庭调查做了介绍。报告引起了与会者热烈的反响。于是在年会期间，费孝通先生趁热打铁，在自己下榻的宾馆的房间举行了一次由与会各单位有关人员自愿参加的吹风会，社会学所参加这次会

议的有潘乃谷和我，其他与会者我记得有天津社会科学院的潘允康（本文所提人物的单位和职务均为当时的情况）、江苏公安专科学校校长姬仰曾、四川社科院社会学所所长詹一之、上海社科院社会学所的薛素珍、北京经济学院的王胜泉和上海高校的与会者等十几个人。他向与会者介绍了东河沿调查的情况，提出在其他城市进行同样调查的建议。他明确指出，参加不参加，由各单位自己决定，姜太公钓鱼，愿者上钩，参加的单位自筹调查经费和差旅费，中国社科院社会学研究所可出具证明文件，承担会议和数据处理费用，以及组织协调与处理数据等技术工作。会议气氛很热烈，与会者都表现出强烈的参与愿望。姬仰曾和詹一之本身就是单位的负责人，所以当即就拍板参加，另一些则表示自己愿意参加，但须回去请示领导后再定。会后，各单位很快就有了正式的回应。最终确定参加的单位有：中国社会科学院社会学研究所、北京经济学院、天津社会科学院社会学研究所、江苏省公安专科学校、四川省社会科学院社会学研究所、成都市社会科学研究所、上海社会科学院社会学研究所、上海复旦大学和上海复旦大学分校等（后来的上海大学）。当时称联合调查组。

联合调查组成立后的第一项工作就是设计用于该项调查的问卷。这一项工作是在北京的东方饭店举行的。与会者以东河沿调查的问卷为蓝本，逐项讨论斟酌，形成了五城市联合调查组使用的统一问卷。调查并没有规定统一的名称，由各参加单位根据实际情况自行确定。各个单位的调查点也是由各个单位根据情况自行选定，而非由事先规定的程序统一抽取的。最后选定并实施的调查地点是：北京宣武区椿树街道东河沿居委会、北京朝阳区团结湖街道团结湖居委会、天津河西区尖山街道红星第二居委会、

上海市徐汇区新乐街道张家弄居委会、上海杨浦区宁国街道双阳居委会、上海虹口区长春街道、南京夫子庙街道四福巷居委会和成都如是庵街居民段。在调查对象的选择方面，除上海虹口区长春街道采用的是街道内的简单随机抽样外，其余的都采用了整群抽样，即调查对象是调查地（抽取的户中）全体已婚妇女。除已经先期完成的天津尖山街道红星居委会和北京东河沿之外，其余调查都是在1983年初实施的。所有调查都是采用调查员入户访谈的方法来填答问卷。

1983年，在全国社会科学规划会上，这一调查被确定为“六五”全国哲学与社会科学的重点项目，项目被冠以“中国城市家庭研究”这一名称，但一般习惯上我们仍叫它五城市调查。项目聘请了雷洁琼先生担任学术指导。费孝通先生则改任另一重点项目，“小城镇研究”的学术指导。最后确定的项目组成员有：中国社会科学院社会学研究所刘英（项目负责人）、沈崇麟；北京经济学院经济系刘金云；天津社会科学院社会学研究所潘允康（项目负责人）、张雅芳；上海社会科学院社会学研究所薛素珍（项目负责人）、杨善华；上海大学文学院社会学系刘炳福、仇立平；上海复旦大学哲学系范伟达；江苏省公安专科学校夏文信、宋践；四川省社会科学院社会学研究所李东山；成都市社会科学研究所吴本雪。那时，早前联合调查组有些人，比如沈关宝，记得在东方饭店讨论问卷的会上，我和他为了家庭的操作定义曾争得脸红耳赤，他在项目组成立时已经离开，随费孝通先生去了小城镇项目组。而复旦大学分校也成为了上海大学的一部分。项目实施过程中，作为学术指导，雷洁琼先生给予了我们极大的支持，她将自己的经验学识倾囊相授。她不仅参加了项目组在北京的全部会议，而且还参加了项目

组的连云港会议和厦门会议。每一次会议都会亲自参与项目组的讨论并提出自己的意见和建议。可以说没有雷洁琼先生的指导就不会有项目的顺利完成。她的“要用资料说话”“论文写作要规范”和“社会学要面向中国的实际”（参见杨善华、王思斌《雷洁琼学术生涯二三事》）的学术研究导向对我今后的研究有很大的影响。五城市调查之后，我参加和主持的各个家庭调查项目，都请她担任学术指导，每一次都得到了她的指导和支持。1996年中日合作进行的城乡家庭调查，九十高龄的雷老还亲赴太仓，出席太仓调查的调查员动员会。不仅如此，雷老也给了东河沿调查基地以无形的支持。在费老因为大家知道的原因离开社会学所之后，东河沿调查基地处于无人问津自生自灭的境地，我出于一种责任和对东河沿的感情，还在那里苦苦支撑。除了平时常去那里看看之外，逢年过节都会自掏腰包和科研处的同事陈婴婴、赵克斌等一起给居委会送去一些可乐、挂历之类的东西。在调查开始或结束时，常常会请居委会的主任们聚餐，联络感情。这时我都会请雷老到场，在这样的场合，雷老与居委会的主任们谈天说地，聊家常，从不说官话套话。居委会主任们感到十分亲切，在他们眼里雷老不只是一位大领导，也是一位可亲可敬的老大姐。一次，在王府大街八面槽翠花楼的聚餐会上，居委会主任们对雷老说，请她以后还要来，雷老风趣地说："当然要来，有说有笑，还有吃。"雷老的话音一落，响起了一片笑声。而正是在这笑声中，东河沿基地得到了雷老无形的支持，东河沿调查基地又延续了十余年。在此期间，我们在那里进行了七城市家庭调查的东河沿调查，接待了“亚太地区家庭未来讨论会”和“中日家庭比较研究”的与会代表。日本家庭研究者访问了东河沿的居民家庭，发现了一种在日本从来没有见过的家庭形式——隔代

家庭。后来担任过英国牛津大学中国研究中心主任的彭轲教授的博士论文中的实地调查，有相当大一部分就是在那里做的。而这一切之所以得以实现，都离不开雷老的无形支持。

在这里，我也要向项目组历次会议的组织和资助者表示自己的谢意和敬意。他们是中国社会科学院社会学研究所的潘博儒、江苏省公安专科学校校长姬仰曾和《家庭》杂志社社长叶先生（时间久远，已经想不起名字）。

三、五城市调查数据的影响

五城市调查的数据全部用计算机处理，这也是费孝通先生倡导的。东河沿调查之初，费孝通先生访美归来，召集调查组开会，谈了自己访美的观感，特别谈到在美国看到用计算机处理调查数据的感受，他用带着浓重的吴江口音的普通话说道："只看见数字（在屏幕上）蹉蹉（飞快）地出来，这些数字不就是文章吗？有了它们还愁写不出文章？"他又一次对我们说"不入虎穴，焉得虎子"，叫调查组用计算机来处理东河沿的数据。当时所里既没有计算机，也没有计算机人才。于是他请了自己的弟弟，当时的华中工学院自动化系仿真实验室的副教授费奇来帮忙，费奇又找了自己的助手刘国民。而我这个鸭子也就上架成了计算机"人才"。在费刘两位先生的指导下，我开始了计算机数据处理的前期工作，学习BASIC语言，编制编码手册，编码和过录（将问卷上的编码登录到一张张专门用于数据录入的专用表格上）。五城市调查虽然是后编码的，但是每一个调查变量的分类都是事先已经确定了的，所以编码手册的编制只需给每一个类别定值，确定所有的变量值在过录表，也就是

将来在数据文件中的位置和字长（即占用的格数）即可。因为字长需事先规定，所以对于数字型变量，视其占据格数的不同，以8和8的组合作为最大的码值。而将9和9的组合定义为缺失值。在此期间，我们所迎来了几位第一届南开社会学专业的毕业生，其中一位叫谢文，被分到了家庭研究室，他很快就投入到项目的工作中。他提出在数据录入之后，再进行一次逻辑检验（跨变量地检查数据存在的错误）。于是我们又一起编制了一份逻辑检验的清单，递交给华中工学院费奇，请他们据此编制逻辑检验的程序。一切准备就绪，我和谢文一起，背着问卷和已经过录好的过录表，赶赴武昌华中工学院，在那里的自动化系仿真实验室的机房，通过他们用扩展BASIC语言编制的录入程序和逻辑检验程序，完成了数据录入、逻辑检验和改错工作，获得了撰写调查报告所需的统计数据。东河沿调查是第一个进行计算机数据处理的调查，以后各地都照此办理，在华中工学院完成数据的录入、逻辑检验和改错工作，并产生将来写调查报告需要的统计数据。当时没有像现在这样供数据处理的软件包，如SPSS、SAS和STATA等。数据处理的程序都是用费奇和刘国明先生，用扩展BASIC编写的程序进行的。

确切地讲，数据录入是分城市分卡进行的。即每一个调查点的数据，分成三卡（三套过录表）按卡一条一条（一行一行）记录输入计算机，生成与三个卡（有关卡的问题可参阅《五城市联合调查组》的问卷）对应的数据文件，存储在一张8吋软盘上。所以五城市调查的原始数据的物理形式是8张8吋软盘，每一个软盘都有三个对应于三个卡的数据文件。数据文件使用的计算机码是CROMCO Z80系列，而非IBM系列的ASCII码。1983年我用所里的科研经费结余，购置了一台由北京计算机研究所和计算机五厂生产的BCM

Ⅲ。从此社会学所和项目组有了自己的计算机。这台计算机也是CROMCO Z80系列的，使用的也是两个8吋软驱。年底，项目组在北京举行《中国城市家庭——五城市家庭调查报告和资料汇编》（山东人民出版社，1984年，济南）一书的编辑和定稿会议时，我们就是用那台机器对五城市调查的数据进行分析，为报告提供所需的各种统计数字和表格。为了实现这一目的，我们首先对东河沿和尖山街的数据，按照后来联合调查组的数据格式，重新进行了编辑和调整，以使数据格式与联合调查组其他调查点一致，产生了格式统一的五城市8个调查点的数据。

1983年6月，广州社会科学研究所的于幼军主动提出用五城市调查的问卷在广州进行一次调查，调查结束后，何颂扬先生携问卷和过录表来京，我们也是用这台计算机处理了广州的数据。这样五城市8个调查点的数据，就扩充成了六城市九个调查点的数据。由于这项工作完成在《中国城市家庭——五城市家庭调查报告和资料汇编》定稿之后，所以该调查的数据和报告并未收入该书。

1985年出版的《中国城市家庭——五城市家庭调查报告和资料汇编》（五城市项目组，山东人民出版社，济南，1985年）一书是项目组的共同成果。责任编辑房喜兰女士是一位资深编辑，独具慧眼看中了这一选题，承担起该书编辑、出版和发行的全部工作。该书出版对国内外的社会学界产生了很大影响，对家庭社会学研究和抽样调查起到了很好的推动作用。1986年我去荷兰求学，学位论文就是使用这本书刊登的资料。我记得我第一次去莱顿大学汉学院去见学校为我推荐的指导教师彭轲（Frank Pieke）的时候，他从办公桌的抽屉里拿出了一本书，对我说，你是研究中国家庭的，一定要看这本书，这本书就是《中国城市家庭——五城市家庭调查报告和

资料汇编》。当时我也很惊奇，没有想到山东人民出版社也将这本书发行到了国外。而实际上我的学位论文使用的最重要的数据都来自这本书。根据分析得到的被调查人分年龄组的家庭结构分布，我论证了中国人一生生活的家庭结构是随年龄变化而变化的，与西方不同，中国的核心家庭，并非婚姻的产物，而是从夫方父系家庭这棵大树生长过程中生长出来的。主干家庭，而非联合家庭，也非核心家庭才是中国传统家庭的基本结构。在各种家庭结构类型构成中，由于传统的中国大多数家庭都是多子女家庭，因此核心家庭所占比例始终都是最高的，并不存在现代化理论所谓的核心化趋势。我的学位论文得到他的高度赞扬，为此他专门给校方致函，对论文做了很好的评价。我在ISS的指导教师，将论文寄给了荷兰家庭研究会会长和跨大学人口研究所（NIDI）所长库耶（Kooye）教授。他也对论文做了较高的评价，认为我的观点很有新意。他建议在此基础上做进一步的研究。因为分组数据，毕竟还是截面数据，如果能用纵向的数据来证明这一结论，便会有更大的说服力。正是根据他的建议，我在以后自己主持的调查中，如七城市家庭调查和国情调查的问卷中都设计包含了一张家庭生命过程表，用来收集被调查人从出生到调查进行时生活的家庭结构类型信息。为了给我今后的研究提供更多的帮助，他邀请我去他乡下的家中详细讨论一下这个问题，他说他家里有一个有关家庭研究书籍和刊物的小型图书馆，可为这个研究提供很多资料。由于在电话中没有听清约会的时间，结果还害得老先生和他的夫人在乡下的火车站白等了好几个小时。那时我已经在准备回国的各项事宜，在离开荷兰之前再也没有机会访问库耶先生的家。后来我的同事折晓叶去荷兰访问，专门去荷兰的乡下他家中，代我探望库耶先生。在完成ISS的学业后，我应邀

到NIDI做为期三个月的访问，利用这三个月的时间，我对这本书汇编部分的统计资料进行了二次分析，完成了一篇中国家庭研究的综述性报告。NIDI的同事对这个研究表现出极大的兴趣，希望今后能一起用原始数据对中国的婚姻家庭做进一步的研究。我现在还记得我的同事巴特雷马当时问我的问题："Where is the raw data?"我现在做的工作就在回答他34年前提出的问题。杨，我希望现在我做的工作，能实现你三十年前想分析的那套资料愿望。如果有机会，我希望能有机会像当年在NIDI那样，我们坐在计算机前一起来分析它。

1990年，我受社会学研究所委托，组织了联合国教科文组织委托的"亚太地区家庭未来讨论会"。这个会对于我来讲，有两大收获。首先，我在会上以《中国城市家庭——五城市家庭调查报告和资料汇编》和人口普查数据为依据，探讨人口要素对家庭结构的影响。我先用五城市分年龄组的家庭结构类型的分布，论证了传统中国家庭结构主干家庭为干、核心家庭为枝的结构形态。而支撑这种形态的人口要素是出生与死亡达到某种相对的平衡，即替代水平不低于2.1。但是人口普查的数据表明，一胎化后中国人口替代水平已经低于2.1，而且呈进一步下降的趋势。这意味着支持主干家庭为干、核心家庭为枝这种结构形态的人口要素已经不复存在，因而家庭结构形态必将随之发生变化，而这一变化，必将对我国社会其他方面产生巨大影响。当时我的报告的评论员是当时日本家庭社会学会的会长森冈清美（森冈先生是我的同事张萍推荐的）。森冈先生不仅对我提出的观点很感兴趣，而且也对我使用的数据源自的调查——五城市调查感兴趣。会后森冈先生对我说，以前对中国的家庭研究不太了解，现在才知道，原来中国在这方面已经做过这样大

规模的家庭调查。他说这是中日双方今后开展合作研究的基础。他表示，回去以后很快就会组织一个老中青三代的研究团队，开展中日家庭的比较研究。森冈先生果然兑现了自己的承诺，次年就率领一个由老中青三代组成的日本家庭社会学家组成的代表团，在北京和杭州两地举行了中日家庭比较研讨会。在杭州，森冈先生的学生，后来的日本家庭社会学会会长，东京都立大学教授石原邦雄来到我的房间，由社会科学院日本研究所的王伟先生做翻译，讨论了开展合作研究的问题，达成了开展中日家庭的比较研究的口头协议，由双方共同讨论决定调查问卷和入户访谈提纲，调查和研究经费由日方申请，中方负责选择调查地、抽样和数据处理，共享调查数据。自此，我和日方团队的合作研究延续了20年。在此期间进行了两次抽样调查、一次深入访谈和多次互访和讨论会。后来在石原先生的推动下，又吸收了韩国家庭研究者，将中日比较研究，发展成中日韩的东亚家庭比较研究。今天活跃在社会学研究和教育第一线的老师，清华大学的王天夫、南京大学的吴愈晓、中央财经大学的范欢欢和中央民族大学的王纪芒等人的学位论文数据都来自这些调查。社会学所的同事陈婴婴和日本研究所的王伟都对这项持续20年的合作研究做出了很大贡献。

另外，我与所长陆学艺达成一个口头协议，一方面，我一定按联合国方面的要求，完成讨论会的全部工作，另一方面，我也希望会议结束后，所里不要将结余的经费挪作他用，而要用它来购置一台计算机。在我按所里的要求完成了讨论会的所有工作后，陆学艺所长也兑现了自己的承诺。于是我用剩余的会议经费2000美金，购置了所里第一台IBM系统的PC。这是一台美国华裔生产的sun286，有了这台机器，我才有可能开展以后的中国百县市调查和

七城市家庭调查。那时五城市调查已经结束多年，并不需要用它处理五城市的调查数据。但是它使我萌发了将五城市调查原来保存的CROMCO Z80系列的ASCII码数据转换成IBM系列的ASCII码数据的想法，而且立即诉诸行动。我请人用这台机器模拟打印机，将从BCMⅢ上读出的CROMCO系列的数据，存入这台机器，完成了将五城市调查数据从CROMCO系统到IBM系统的转换。现在想起来，这次转换是很及时的，因为不久之后机器的更新换代越来越快，这两台机器很快就不敷使用，为功能更为强大的新微机取代而被弃之不用。这就是说，如果当时未能及时将这些CROMCO系列的文件转成IBM系列的文件，那么今天重新整理这批数据就会大费周章了，真是万幸。

正是在那一年，原来五城市项目的刘金云先生从密歇根大学回国探亲，给我带来了一套DOS版的SYSTAT，这是我使用的第一套数据分析软件。之前我曾经试着在SUN上安装DOS版的SPSS，但是由于SUN的系统与IBM并不完全兼容，DOS版的SPSS无法在SUN系统中正常运行。而SYSTAT则可以在SUN系统正常运行。我立即用它生成了一套SYSTAT格式五城市调查数据，并与项目组原来的成员李东山合作编辑了一本《中国城市家庭——五城市家庭调查双变量和三变量资料汇编》（社会科学文献出版社，1991年5月，北京）。该书是对1984年出版的那本中国城市家庭资料汇编部分的补充，书中收录的都是双变量和三变量的交互表。

四、五城市调查数据再次整理

2019年，我终于完成与重庆大学出版社合约的最后两本书，

《社会科学研究方法百科全书》和《测量、设计和分析》（待出版）的翻译，休息了一段时间后开始着手整理几十年来自己参与和主持的十几个抽样调查的数据。1992年在与芝加哥大学和杜克大学合作进行《中国城乡居民家庭生活调查》之后，由于合作者芝加哥大学白威廉的推荐，我开始使用STATA。我记得当时白威廉先生给我寄来的第一套STATA是2.0，还是DOS版的。而在开始做这个工作的时候，最新的版本已经是STATA16了，现在已经是STATA17了。STATA14之后的各个版本，除了数据编辑、分析和绘图功能有很大的改进之外，系统采用了新的UNICODE码，因此采用STATA14以前各种版本生成的系统文件的中文标签，在STATA14之后的各个版本中都变成了乱码，无法辨认。而当时我手头有的STATA是12，还是10年前与香港理工大学合作进行手机调查时，理工大学的罗沛霖先生受我影响从SPSS改用STATA，购买赠送给我的。换言之，我对历年调查数据的编辑整理都是用STATA16进行的。当然五城市数据也不例外。编辑整理的过程是，先分调查点将对应于原始调查三个卡的ASCII码数据转换成STATA数据文件，然后分调查点将3个STATA数据文件合并成一个包括三个卡的全部数据，最后将9个数据合并成一个数据——五城市调查数据（wcsdcdata.dta）。借此机会我也对数据做了最后的编辑整理：给数据的变量和变量值贴上了标签。这个工作年轻人做比较合适，老头子，老眼昏花了贴起来有点费劲，但是总算干成了。在这个过程中再次仔细读了全部数据并做了记录。

最终的数据文件是一个多重记录类型的数据文件，有A、B、C三种记录类型，其中A类记录是单条记录类型的数据，记录了被调查人本人的情况。本次调查共有被调查人5748（5075户）其中北

京东河沿701（669户）、北京团结湖（505户）、天津尖山街（583户）、上海张家弄739（646户）、上海长春街791（640户）、上海双阳路775（719户）、南京四福巷629（530户）、成都如是庵401（349户）和广州某居委会528（528户）。由于本次调查的被调查人是样本户中的全部已婚妇女，不少样本户都有一个以上已婚妇女，所以被调查人数会多于样本户数。为便于大家了解数据的结构，我在被调查人数之外，也统计了样本户数，把它们列在了被调查人数后面的括弧中。现在看来，这样的设计叠床架屋，不仅有重复之嫌，也给数据处理增加了麻烦。户被调查人数如下表所示：

回答人数	1	2	3	4	5	合计	调查户数
北京东河沿	460	106	7	2	0	575	701
北京团结湖	454	48	3	0	0	505	559
天津尖山街	545	36	1	0	1	583	625
上海张家弄	556	87	3	0	0	646	739
上海长春街	503	126	9	1	1	640	791
上海双阳路	665	52	2	0	0	719	775
南京四福巷	444	74	11	1	0	530	629
成都如是庵	299	48	2	0	0	349	401
广州某居委会	528	0	0	0	0	528	528
合计	4454	577	38	4	2	5075	5748

B类记录属于多条记录类型数据，记录了被调查人的家庭成员的情况，每一个家庭成员都有一条单独的记录，共有22122条（个家庭成员），说明本次调查涉及的被调查人的家庭成员共有22122人。其中北京东河沿2289、北京团结湖2312、天津尖山街2408、上海张家弄3031、上海长春街2616、上海双阳路3087、南京四福

巷2742、成都如是庵1278、广州某居委会2449。两类记录的拼接结果显示，两类数据的ID完全对应。与B类记录一样，C类记录也是多条记录类型的记录，记录了被调查人的生育情况，涉及被调查人的全部孕育过程，即每一个孩子的怀孕、生育和抚育情况。有多少个孩子就有多少条记录。全部记录共为25353条，除去天津缺失的529条（因为最早开始的天津尖山街调查并没有生育部分的内容），共有有效记录24824条（个孕育的孩子），其中北京东河沿2647、北京团结湖2429、天津尖山街缺失、上海张家弄3452、上海长春街3165、上海双阳路3629、南京四福巷3174、成都如是庵1618、广州某居委会2655。与A和B类数据拼接也正确无误。由于这一套数据不仅时间跨度很大，而且中间经过多次转换。所以在拼接之前我十分担心会出现意想不到的情况，但结果却证明我的担心是多余的。在完成数据拼接之后我又对数据文件所含125个变量一一做了检查。对超界和异常值做了处理。拼接和检查结果说明，原数据清洁整齐，质量很好。至此，或许我可以向费雷二老，和当初一起熬夜设计、挨家挨户登门调查和彻夜不眠处理数据的诸位同道交差了。至于费老关心的“虎子”擒得否？请社会学界同仁指正。

五城市调查最后编辑整理的数据变量清单如下：

wcsdcdata.dta数据变量清单

变量	变量值标签	变量名标签
id		个案编号
jwh	jwh	居委会
fm		户内回答人编号
a1		本人出生年代
a2		本人出生月份

续表

变量	变量值标签	变量名标签
a3		本人出生日
a4		丈夫出生年代
a5		丈夫出生月份
a6		丈夫出生日
a7	edu	本人文化程度
a8	edu	丈夫文化程度
a9	bplace	本人出生地
a10		本人来本市时间
a11	bplace	丈夫出生地
a12		丈夫来本市时间
a13		结婚年代
a14		结婚月份
a15	job	本人结婚时有无工作
a16	indus	本人工作的行业
a17	unit	本人工作单位性质
a18	occu	本人职业
a19	job	丈夫婚时有无工作
a20	indus	丈夫单位的行业
a21	unit	丈夫单位性质
a22	occu	丈夫职业
a23	select	认识途径
a24	dis	婚姻距离
a25		婚时娘家人口
a26		婚时娘家代数
a27		婚时娘家夫妻对数
a28	fstruc	娘家家庭结构
a29	whn	娘家家庭费氏结构 （注：在同一个锅吃饭的人）
a30		婚时婆家人口
a31		婚时婆家代数

续表

变量	变量值标签	变量名标签
a32		婚时婆家夫妻对数
a33	fstruc	婚时婆家家庭结构
a34	whn	婚时婆家家庭费氏结构
a35	job	婚时公公有无工作
a36	indus	婚时公公工作的行业
a37	job	婚时婆婆有无工作
a38	indus	婚时婆婆工作的行业
a39	job	婚时父亲有无工作
a40	indus	婚时父亲工作的行业
a41	job	婚时母亲有无工作
a42	indus	婚时母亲工作的行业
a43	comp	婆家娘家经济比较
a44	yn	有无订婚
a45	yn	是否有登记
a46		婚姻登记 年代
a47		婚姻登记 月份
a48		婚礼年代
a49		婚礼月份
a50	wtype	婚礼形式
a51	lsite	婚后居处
a52	hown	住房所有权
a53		房屋租金(角)
a54		住房间数
a55		住房面积
a56	hhave	有无自建住房
a57		自建住房间数
a58		自建住房面积
a59	htype	住房形式
a60	chave	厨房情况
a61	chave	厕所情况

续表

变量	变量值标签	变量名标签
a62	yn2	是否负担不在一起生活的人
a63		负担人数
a64	rela	被负担人的关系
a65		负担金额
a66	hhave	是否有不在一起生活的人资助
a67		资助人数
a68	rela	资助人关系
a69		资助金额
a70		目前家庭 人口数
a71		目前家庭 代数
a72		目前家庭 夫妻对数
a73	fstruc	目前家庭 结构
a74	whn	目前家庭费氏结构
a75	hsta	目前丈夫 情况
a76		丈夫离婚 或去世年代
a77		怀孕次数
a78		生育次数
a79		生育男孩数
a80		生育女孩数
a81	yn2	是否避孕
a82	yn2	是否绝育
a83		初经年龄
a84	yn2	是否停经
a85		停经年龄
b1	b1valu	与调查对象关系
b2	b2val	性别
b3		出生年代
b4		出生月份
b5	b5val	文化程度
b6	b6val	工作单位的行业

续表

变量	变量值标签	变量名标签
b7	b7val	
b8	b8val	职业
b9		基本工资
b10		月奖金
b11		福利及其他
b12	b12val	是否在家吃饭
b13	b12val	是否在家住
b14	b14val	户口所在地
c1		怀孕顺序号
c2	c2val	子女性别
c3	c3val	流产、死亡情况
c4		流产、死亡年代
c5		流产、死亡月份
c6		出生年代
c7		出生月份 月份
c8		喂奶月数
c9	edu	文化程度
c10		孩子参加 工作年代
c11	c11val	孩子工作单位行业
c12	b7val	孩子工作 单位性质
c13	c13val	孩子职业
c14		孩子结婚年代
c15	lsite	婚后居处
c16	c16val	0–3岁谁照看
c17	c16val	4–6岁谁照看
c18		现有男孩数
c19		现有女孩数
c0		子女编号

吴肃然，孔天慧．从 KKV 到 TTC：近 30 年美国最有影响力的两本质性研究方法论著作[M/OL]//赵联飞，赵锋．社会研究方法评论：第 2 卷．重庆：重庆大学出版社．

从 KKV 到 TTC：近 30 年美国最有影响力的两本质性研究方法论著作

吴肃然　孔天慧[①]

摘要：KKV 和 TTC 是美国社会科学界近 30 年来最有影响力的两本质性研究方法论著作。KKV 尝试将质性研究统合到主流量化研究的框架下。这一工作虽然未获得质性研究者的认可，却意外得到了量化研究者的极高评价。TTC 不同意 KKV 的观点，主张量化研究和质性研究是两种不同的文化。这一论证得到了许多质性研究者的高度赞扬，却未能赢得量化研究者的认同。本文认为，作为当今中国的社会学研究者，在看待从 KKV 到 TTC 的方法论历史时，我们需要打破常规的理解，认识到 KKV 和 TTC 是同一种社会科学哲学思潮的两个分支，二者的共性潜藏着美国主流社会科学的局限，二者的分歧则蕴含着丰富的知识意义。

关键词：KKV；TTC；方法论；质性研究

①作者简介：吴肃然，哈尔滨工程大学社会学系教授，博士生导师，主要研究方向为社会研究方法论。联系方式：wusuran@hrbeu.edu.cn；孔天慧，哈尔滨工程大学社会学系本科生，主要研究方向为定性研究方法。联系方式：kongtianhui@hrbeu.edu.cn。"

Abstract: KKV and TTC are two of the most influential methodological works on qualitative research in the American social sciences in the past 30 years. KKV attempts to integrate qualitative research into the framework of mainstream quantitative research. While this work has got high praise from quantitative researchers, it has not been accepted by qualitative researchers. TTC disagrees with KKV's argument and it proposes that quantitative and qualitative research are two different cultures. While this argument has been highly praised by many qualitative researchers, it has failed to win the consent of quantitative researchers. This paper argues that, as sociological researchers in China, when looking at the methodological history from KKV to TTC, we need to recognize that KKV and TTC are two branches of the same philosophy of social science, and that their commonality reflects the limitations of mainstream American social science, while their divergence implicates a lot.

Key words: KKV; TTC; Methodology; Qualitative Research

美国社会科学界近30年来最有影响力的方法论著作有两本：其一是加里·金（Gary King）、罗伯特·基欧汉（Robert O. Keohane）和悉尼·维巴（Sidney Verba）三人所写的《社会科学的研究设计：质性研究中的科学推论》（简称KKV）；另一本是加里·格尔茨（Gary Goertz）和詹姆斯·马奥尼（James Mahoney）所写的《两种传承：社会科学中的定性与定量研究》（简称TTC）[1]。KKV初版于1994年，该书出版后得到了美国社会科学界的广泛好评，被人们誉为“所有社科研究者的必读”“社会科学研究方法的圣

[1]KKV是三位作者的姓名简写，这一称呼已被广为接受；TTC是书名的简写，系本文作者的用法。

经”。TTC出版于2012年，该书对KKV的观点进行了全面系统的挑战，近几年来也备受学界热议。

一、KKV：用量化研究的逻辑统合质性研究

KKV一书的三位作者都是美国政治学界的重要学者，第一作者加里·金最为人们所熟知，他是哈佛大学的“大学教授”（university professor）、美国科学院院士。KKV被誉为“哈佛三大重量级学者思辨结晶”，是哈佛、哥伦比亚、华盛顿、普林斯顿等大学方法论课程的指定教材，是“近二十年来最重要的社会科学方法论经典”[①]。该书代表着秉持主流量化方法（mainstream quantitative methods）的社会科学家对质性研究所做的批判和改造（Collier，Seawright & Munck，2010）。这一批判的基础立论并不新鲜：科学是对外在世界中客观事实的探寻，科学要遵循明确和公开的步骤和方法，科学研究不能以偏概全，科学表述不应使用散文式的语言，科学应当数字化，不一而足。在100多年来有关社会学方法论的论辩中，这些观点早已被反复地主张过、批判过，但人们从来没有达成共识。在当代社会科学的论域中，这些观点依旧被一些学者拥护而被另一些学者批判。KKV一书的核心目标则是要在技术层面，用主流量化方法的框架来重新打造质性研究，将拉扎斯菲尔德等人早就主张的方法论逻辑贯彻到底（McKeown，1999）。三位作者认为，质性研究者所提出的“质性研究的独特性”等主张以及相应的方法论问题，都可以在主流量化方法的框架下化解（King，Keohane & Verba，1994；

①见KKV一书台湾译本封底简介：盖瑞·金、罗伯特·基欧汉、西德尼·维巴，2012，《好研究如何设计：用量化逻辑做质化研究》，盛智明、韩佳译，台北：群学出版社。

3–9）。

KKV一书的目标读者是从事质性研究的社会科学家。作者试图说服他们：所有的社会研究都可以统合到主流量化研究的框架下。具体来讲，就是基于一种反事实的因果定义，利用以回归分析为主的技术来实现因果推断。在KKV中，三位作者把“选择性偏误”（selection bias）界定为质性研究的核心方法论问题，并使用多种操作性概念来化解质性研究中样本非随机性的问题。这一做法深深赢得了许多社会科学研究者的认可。因为在他们看来，相对于量化研究来说，质性研究并没有什么特殊性，而KKV的论证则将这一观点细化和系统化了。正如谢宇对KKV的评价：

过去，甚至包括现在，无论是在政治学还是社会学，无论是在北美还是在中国，定性和定量取向的学者常常站在不同的阵营，对彼此的工作互不了解，对孰优孰劣争执不休，在各自研究风格的发展上渐行渐远。但这本书就明确整合了这两种方法的共性，即两者都是“系统和科学的工作”，目的都是要提供“有效的描述性或因果性推论”。而且，此书进一步提出，如果定性研究者要作出更为普遍、逻辑关系更清晰、更确证可靠的推论，如果又能突破数据收集的限制，就应该使用定量研究的科学推论原则和建模方法。（金、基欧汉、维巴，2014：中文版序V）

不难想象，对于绝大多数从事质性研究的社会科学家来说，KKV的上述观点既武断又自大。将“选择性偏误”界定为质性研究的核心方法论问题，这种出发点在质性研究者看来本身就是戴着有色眼镜的体现。从论证形式上看，KKV要将所有的质性研究都化约到回归分析的框架下的做法更是不可理喻的。在KKV出版后的十多年里，一些研究者结合自己的研究经验，对KKV的观点进行了

多方位的回应和批判（Brady & Collier，2010）。有的批判甚至非常激烈，比如罗切斯特大学的政治学教授詹姆斯·约翰逊（James Johnson）就很不客气地指出：KKV偷偷搬运的实证主义给美国政治学研究带来了很不好的影响，它所倡导的方法论统合适得其反，给美国政治学界带来了相互的猜疑、指责和误解（Johnson，2006）。国内著名政治学家、社会科学哲学家唐世平也对KKV做出了不留情面的评价：KKV是一部不幸的作品，三位作者不仅不懂定性研究，也没有提及任何已有的对定量方法的深刻批评，KKV一书对方法论的理解有严重偏差（唐世平，2015）。

二、TTC的理论基础：QCA方法

虽然绝大多数质性研究者不赞同或无视KKV对于质性研究方法论所做的批判与整合，但也有一些人对KKV的观点给予了充分的关注与回应，TTC的两位作者加里·格尔茨和詹姆斯·马奥尼便是如此。格尔茨和马奥尼长期从事社会科学的质性研究，在KKV出版后，他们发表了多篇论文，对KKV进行回应和批判。2012年，两人共同出版了TTC一书，将长期的方法论思考集之大成，系统阐述了自己对于KKV“统一方法论主张”的异议，捍卫质性研究的独特性。

如同KKV依托于回归分析，TTC的写作也基于一套经验分析技术，这套技术被称为定性比较分析（Qualitative Comparative Analysis，QCA）。QCA由美国社会学家查尔斯·拉金（Charles Ragin）开创，它是一套服务于定性分析的研究技术，是质性研究中极其少见的、有着明确逻辑步骤的方法论框架，被视为近几十年来西方社会

科学研究中少有的具有实际意义的方法论创新（马克斯、里候科斯、拉金，2015）。1984年，拉金在《美国社会学评论》发表了一篇讨论歧视的论文，开创性地将布尔代数法介绍给美国社会科学界（Ragin，Mayer & Drass，1984）。1987年，他出版了《比较方法：超越定性与定量》（*The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*）一书，系统介绍了布尔代数法的原则及其在社会科学研究中的前景。在该书前言中，拉金写道：

> 我创立和发展一个质性整体比较研究技术的兴趣，来自我作为一名比较社会学家所体会到的挫败感。如同今天绝大多数的美国社会学家一样，我所接受的训练是尽可能地使用多变量统计技术。然而我时常发现，对于我感兴趣的一些问题来说，这些技术根本不适用……起初我并没有试图开发适用于这些问题的新技术，而是尽可能地用传统统计方法来处理。这么做的远不止我一人。现如今，社会科学家们总是在大数据的支持下，用多变量统计方法来回答所有的研究问题。我们常常能够看到，为了能够使用某些技术，社会科学家就对自己的问题进行削足适履的操作。（Ragin，1987：vii）

对于关注方法论争议的社会科学研究者来说，拉金的这番表述深入人心，有非常多的质性研究者都指出过定量研究技术的片面性以及它对研究问题所产生的逆向影响。不同于大多数人的是，拉金没有停留在理念批判上，而是开创了一种开放的、具有复制性的、有明确逻辑的质性研究方法，这套方法依托的数学基础不是主流量化研究中的基础统计学，而是布尔代数（Boolean algebra）。布尔代数是英国数学家布尔在1854年提出的一种逻辑代数方法，但是直到罗素与怀特海的《数学原理》出版之前，布尔代数没有得到数学界与逻辑学界的认可。在罗素等人借鉴了布尔的观点之后，布尔代

数方法与分析哲学运动日渐结合，从而对逻辑哲学产生了深远影响，在罗素和早期维特根斯坦的数学哲学那里，这一点体现得非常明显。布尔代数在20世纪的影响不仅仅发生在哲学领域，如果我们对近几十年来的通信技术和计算机科学有所了解，就能很清楚地看到以布尔代数为标准的基本逻辑架构。

布尔代数与回归分析的本质区别可以从一个简单的方程中看出来：

$$Y = A + B + C*D$$

在主流量化技术的视角下，这会被看成一个回归方程，它包含四个自变量A、B、C、D和一个因变量Y，在变量C和D之间存在着交互效应。变量A、B、C、D都是定距变量，由它们的取值可以计算出对应的Y值。当然，这个回归方程有些特殊，它的回归系数都是1，同时方程中缺少误差项。在布尔代数的视角下，此方程的含义有很大差别。变量ABCDY全部是二分变量，它们的取值只有两个，即0或1。方程中的“+”意味着逻辑关系“或”，“*”则意味着逻辑关系“与”。在通常的回归分析中，如果变量是二分变量，则需要通过取对数的方法将其转化为定距变量，其理由是二分变量是级别较“低”的变量，不宜采用回归分析的方法来处理。而布尔代数直接处理的就是此类变量，它关注的是一个变量的有无，即某概念的性质，这种“性质”正是质性研究的焦点。

根据上述模型，Y只有两个结果：发生或不发生，ABCD的取值即是对应的发生条件。如果Y=1意味着某国家发生革命，那么在共计13种情况下，国家会发生革命。当A条件满足时，肯定会出现革命；当B条件满足时，肯定会出现革命；当C和D同时满足时，也肯定会出现革命。不难看出，这种逻辑思维与历史社会学中的比

较分析十分接近，实际上，拉金正是在研读巴林顿·摩尔的著作时获得了QCA的方法论思路。

布尔代数方法的发问方式和主流量化研究有比较明显的差异，后者不可能回答"在哪些条件下会发生国家革命"此类问题，而总是去识别某个具体条件与国家革命之间的因果关联。拉金认为，采用布尔代数法有助于社会学家摆脱量化研究框架的束缚，从而能够回答那些真正困扰着研究者的理论问题。

拉金的《比较方法》一书引起了学术界的广泛关注，以政治社会学领域的学者为代表，学者们陆续学习、应用该方法，相应的研究成果有相当高的比重都发表在美国最权威的学术期刊上（马克斯、里候科斯、拉金，2015）。在这之后，美国政治学界也逐渐开始关注拉金的工作，越来越多的政治学研究都采用QCA方法来设计。到了2010年以后，定性比较分析已经成为美国政治学研究的主要研究方法之一。相比起来，美国社会学界对QCA的接受度反而较低；许多社会学家还固守在主流量化思维中，这是一个有趣和令人遗憾的现象。

在拉金最初的QCA框架中，变量只有0或1两个取值，这被称为清晰集QCA（Crisp-set QCA）。后来陆续又有学者发展出了多值QCA（Multi-value QCA）和模糊集QCA（Fuzzy-set QCA），前者指的是一个变量可以有0、1、2等多个取值，后者指的是一个变量可以根据自己的概念隶属度获得［0，1］之间的一个取值。由于QCA没有考虑到变量之间的时间关系，又陆续有学者通过将时间变量引入QCA，设计出了时间QCA（Temporal QCA）和时序QCA（Time-series QCA）。关于QCA的总体发展史，中国学界已有不少文献进行讨论（释启鹏，2016；夏鑫等，2014；李蔚，何海兵，2015；何

俊志，2013），这些讨论的作者也多具有政治学背景。

三、TTC对KKV的批判

在发展QCA的学者中，有一些人就它的方法论意涵进行了深入的讨论，加里·格尔茨和詹姆斯·马奥尼的工作就是代表。在2012年出版的名著TTC中，格尔茨和马奥尼提出：

> 通过强调定性研究与定量研究的差异，我们表达了与KKV一书相反的立场。他们的著名论断是："定量传统与定性传统只有风格上的差异，这些差异在方法论上和实质意义上都是无关紧要的。"他们认为，两种传统共享着同一种推断逻辑，这种逻辑大体上可以用统计分析的规范加以总结。他们依赖一些表面特征来界定两种传统，尤其是对于数字和词汇的使用。我们拒绝如下假定，即定量研究和定性研究可以由单一的建立在统计分析之上的推断逻辑所引导……KKV声称："定量研究与定性研究之间并不存在着孰优孰劣的问题"，然而之所以得到这个结论，仅仅是因为他们觉得只有当统计分析不可行的时候，才不得不使用定性方法。与之相反，我们认为，定量和定性技术各自适用于不同的研究任务，它们是为了达成不同的研究目标而被设计的。（Goertz & Mahoney，2012：2-3）

可以看出，TTC的方法论立场非常鲜明，即定量研究和定性研究所依托的是两种文化，而KKV的主张代表着"主流量化方法"的文化霸权。其实这种立场在质性研究者那里很常见，但是与大多数质性研究者所发出的"弱者的反抗"不同，TTC明确、丰富、清晰地论证了这种方法论立场且对主流量化思维做出了有力反驳。这种论证和反驳正是依赖QCA这一工具达成的。

TTC的主要方法论观点和对KKV的反驳大致从以下方面展开。

第一，因果关系的不对称性。在主流量化方法的文化中，因果关系通常都是对称的，即如果存在因果关系，那么不论自变量是增加还是减少，自变量的变化必然都会带来因变量的变化。TTC用QCA所内含的集合论思维指出，之所以主流量化研究都只处理对称性的因果关系，是因为它们通常意识不到还存在着非对称的因果关系，并且回归分析也只能处理对称关系。从逻辑学的语言上讲，所谓对称关系，即意味着自变量X是因变量Y的充分必要条件。然而在社会现象中包含着大量的充分非必要条件和必要非充分条件，这些都是非对称关系，用回归分析的方法来考察这些关系是错位的，也无法揭示现象背后的因果本质。以美国一流大学社会学系的录取过程为例，计算列联表相关系数可以得出在申请者的GRE数学分数与录取结果之间存在着因果关系。然而，在模糊集的视角下，上述因果关系实际上是非对称的。GRE的数学分数是录取的必要条件而非充分条件。通过这种QCA的分析，我们才能真正识别出美国一流大学的社会学系在录取学生时的标准（Goertz & Mahoney，2012：66-68）。

第二，因果关系的情境性。主流量化方法虽然也强调因果关系的情境性，但这种情境都是以限定词出现的，比如研究者会在论文题目中标定自己所做研究的时间跨度、地点范围或研究对象的特征。这种做法并没有真正处理情境性问题，而只是将情境视为研究的外在条件，将其预先排除出去。主流量化方法认为，在进行了这种情境限定后，自己所考察的X与Y之间就必然存在着唯一确定的因果方向和因果强度，研究者的任务就是把这个数字找出来。在QCA的视野下，这种处理也是不妥的。以从事中国研究的社会分层

学者为例，他们常常在研究中考察党员身份与收入、阶层和干部提拔机会之间的因果关系，并得到结论：党员身份或有助于或无助于各种机会获得。这种考察将中国的“党员机会获得问题”塞到了帕森斯的模式变量的理论思维和回归分析的技术工具中，从而得到了所谓“简化”和“确定”的结论。熟悉中国现实的人们都知道，虽然党员身份对于升迁很重要，但中国在干部提拔中存在着“无知少女”的现象，即拥有无党派人士、知识分子、少数民族和女性四重身份的人比较容易在干部选拔中脱颖而出。此时我们能够看到，在这个变量组合中，非党员身份对于干部提拔起着非常重要的正向作用。由此我们能够发现，在党员身份与干部提拔机会之间并不存在着唯一确定的因果关系，在不同的情境下，党员身份这个自变量所发挥的作用有可能是相反的。这才是经验现象背后的社会事实。然而到了回归分析的框架中这种事实就被隐藏了，研究者或是通过计算平均效应从而得出“因果关系不存在”的结论，或是把少数现象当成误差予以排除。

第三，因果关系的同因果性。从不对称性和情境性可以看出，在QCA的思维中，一个社会现象有可能存在着多个原因路径，比如A+B是Y的充分条件，同时B*C也可能是Y的充分条件。这种情况在社会现象中屡见不鲜，然而主流量化方法也忽视了这一点。主流量化方法在考察因果关系时所依据的是反事实定义，即考察在引入A和不引入A的两种情况中，Y发生了何种变化，以此计算A与Y之间的因果关系。如果存在着同因果性的情况，那么原本存在的因果关系就很容易被误排除。

第四，大数量（Large N）的盲见。主流量化研究认为质性研究所做的小数量（Small N）研究存在代表性问题，而自己依据中心极

限定理等一系列数量统计理论为基础的技术为研究提供了可靠的科学保证。TTC从多个角度批判回应了这些观点。主流量化技术依托正态分布来计算误差，在该技术体系中，最重要的是均值，误差最小的是处于分布最中间的采样点，那些极限值是误差最大的。这种处理其实有悖于许多社会测量的实际情况，比如当人们对威权国家和民主国家打分时，那些得分最高和最低的往往是测量最准确的（如朝鲜），那些位于中间值的（比如一些南美国家）国家反而方差最大。另外，在测量中存在着"无意义变异原理"，这也是量化研究者常忽视的问题。按照QCA的原则，测量应当依据语义学原理，尽管姚明比科比的个子高26公分，但是就"高个子"这个概念来说，二者的测量值应当是一样的。与之相比，在160—186公分这个区间，变量的每个变化都会带来概念隶属度的很大变化，研究者需要根据文化和语义来进行操作化。再次，主流量化研究对变量进行对数转换的做法，常常扭曲了真实的社会现象，也属于让研究对象服务于技术的错误做法（Goertz & Mahoney，2012：150-160）。

TTC一书的出版在美国社会科学方法论领域引起了很大反响，正如拉金所评价的："定性研究一直被描绘成定量研究的庶子……TTC则终结了这种谬论。"（Goertz & Mahoney，2012：封底）许多美国政治学家和管理学家都曾讨论过该书重要的方法论意义，然而与KKV一样，这些真正重视TTC的学者，大都本来就是从事质性研究的，或至少对质性研究方法论抱有同情态度。我们很难见到有哪位KKV的信徒在读了TTC后发生了方法论的转变，更为可能的是，他们根本就不会关注或细读这本书。

四、从KKV到TTC：知识效果与延伸评价

综观KKV和TTC两本著作，我们能看到它们所引发的一些实际知识效果，也可以从方法论角度做出一些延伸性的评价。

第一，KKV和TTC的知识效果体现了一个常见的、稳固的知识社会学现象，即在不同知识阵营的对抗中，对立双方的论辩都无法说服对方，无法实现与对方的沟通理解，反而强化了己方的内部认同。

KKV虽然以质性研究者为自己的目标读者，但是从绝大多数质性研究者的角度看，KKV是一部失败的方法论著作，或可称其为"伪质性研究方法论"。然而十分吊诡的是，KKV一书获得了非常多的关注和好评。之所以会出现这种现象，其原因在于，许多从事量化研究的学者认为，KKV系统完整地表述了主流量化研究的方法论立场和技术细节，认为该书系统、清晰、细致地澄清了主流量化研究的逻辑。于是，在质性研究者对此书毫不感冒的同时，量化研究者却对其进行了大量的学习、转述和引用，最终让这本书成为了主流量化研究的"圣经"。在该书出版后的20多年中，这种实际呈现的知识效果是相当意外的。

与KKV咄咄逼人的"进攻"姿态不同，作为方法论场域中相对处于守势的质性研究，TTC对KKV的批判采取了一种较为温和、隐晦的策略，但这种策略也不成功。尽管TTC指出了主流量化研究自身的许多盲见，但两位作者很少直接展开批判，而主要从捍卫质性研究独特性的角度来展开论证，其对量化研究的态度更多是以副文本的隐晦方式展开的。从这本书的英文书名"A Tale of Two Cul-

tures”就能看出二人的良苦用心，他们试图借用“双城记”（A Tale of Two Cities）的典故放低姿态，呼唤宽容和理解，让主流量化研究能够认识到自己的局限性，为质性研究争取正当性。在英文原著的封面配图上，普林斯顿大学出版社特意使用了著名画家杜尚的名作《下楼梯的裸女》，这幅画背后有着丰富的艺术史意涵，隐晦地表达了对于量化研究的文化批判。但即便如此，也没收到良好的沟通效果。TTC出版以来，几乎见不到任何有影响力的量化研究者对其予以真正的关注和理解，与KKV一样，它也几乎被目标读者忽视了。

第二，从方法论的学术逻辑上看，TTC对于KKV的确有着很大超越，对质性研究方法论也做出了实质贡献。它澄清了质性研究的某些隐含逻辑，揭示了质性研究在方法论方面的某些独特性，有力地反驳了KKV和主流量化研究对于质性研究的批判。对于许多质性研究者来说，上述隐含逻辑往往是一些轮廓不清晰但又比较牢固的观点，TTC将其予以澄清，这种工作有着非常重要的方法论价值。然而，遗憾的是，除QCA的几位核心学者外，其他许多学者往往未能清楚认识到这一点。对于许多熟练运用QCA的学者来讲，TTC的方法论价值都是笼统和模糊的，他们大都将TTC视为对KKV的方法论翻转，这种看法就过于简单了。

以《新教伦理与资本主义精神》为例，韦伯在书中论证了新教伦理与资本主义精神之间的因果关系，对于此因果关系的具体含义，后来的学者展开了诸多论辩。有人举出非新教国家的资本主义发展作为反例来批评韦伯，有人采用大数据的多变量分析，发现新教伦理与资本主义精神两个变量间没有那么强的因果联系，认为韦伯的研究充其量只是提出一个假设，而该假设是错的。众所周知，

韦伯自己也就此问题做出过方法论的澄清，在他看来，社会科学特别是理解社会学所发现的因果关系只是一种“选择性亲和”（elective affinity），它和物理学中的因果关系是不太一样的。围绕着“选择性亲和”这一康德式的辩解，社会科学界又展开了经久的辩论（王水雄，2001）。然而，如果借助QCA对因果关系所做的澄清，我们对于《新教》一书就有可能获得新的理解，从而拨开长久以来笼罩于此书的方法论迷雾。

除因果关系问题以外，从事质性研究的学者最常遇到的责难就是个案研究的代表性问题。在定量研究“Large N”的整体实证逻辑框架下，无论质性研究者做出什么样的论证，都显得很苍白。相比来说，TTC针对个案所提出的方法论辩护就比较有说服力：个案研究是过程追踪的主要方式，在此过程中研究者可以进行两种重要的检验，即箍筛检验和确凿证据检验（格尔茨、马奥尼，2016：106-109）。这种观点就改变了以往方法论辩护就理论谈理论的做法，打通了方法论的正当性与微观研究技术之间的联结。

此外，QCA的布尔代数赋值法还挑战了主流量化研究的方法霸权，真正在技术层面展现了质性研究的独特价值，即概念中心的立场。在主流量化研究那里，变量被分为不同层次，即美国心理物理学家史蒂文斯（Stevens，1946）提出的四种测量层次：定类、定序、定距和定比，依次往后，适用的数学运算就愈加复杂。这种测量层次的划分与统计学紧密联系，在这种统计观念下，使用分类法的质性研究就成为最低测量层次的应用者，其“科学地位”可见一斑。但是QCA完全推翻了这种成见，它利用“无意义变异原理”，将概念的意义而非某种操作化结果置于研究的中心，这种做法已经把社会学方法论争论中的一些关键点，如研究者主体的作用、社会

文化的背景建构等因素带入了QCA的框架中，而且具有很强的逻辑性和实践性，这是许多质性研究传统都没能做到的。

第三，TTC虽然对质性研究方法论做出了重要贡献，但它也有着明显的局限性。长期以来，社会科学的质性研究并不像量化研究一样有着较为统一的研究规范和程序，人们常能见到的质性研究传统有民族志、理解社会学、结构主义、现象学社会学、常人方法论、叙事分析、话语分析等，这些传统与QCA都有着完全不同的哲学源头，在理论关怀和研究程序上有着显著差异。其实，格尔茨与马奥尼已经明确意识到这一点，他们在TTC一书的引论中说道："我们对因果推断的关注有一个后果，即定性范式内部的一些重要流派将不在分析之列。特别是，在我们的两种传统论证中并没有描绘诠释性研究。这些研究路数通常不关注因果分析；它们更专注于其他研究目标，如阐释行为的意义或者批判权力的运用。诠释性传统有自己的主导规范和实践，基本上不同于本书所研究的定量和定性范式。人们当然可以另撰专著，专门考察诠释性传承与我们所讨论的'因果推断'传承在哪些方面可作对比。这样一部书会公开社会科学的各个部分内部存在的认识论和本体论的根本冲突。然而在本书中，我们关注那些在认识论和本体论的一些根本议题上达成一致——包括认为因果分析在理解社会世界方面居于核心——的学者。"（格尔茨、马奥尼，2016：4—5）换言之，TTC一书讨论的只是质性研究的一个子传统，此子传统与其他的质性研究传统有着比较远的方法论距离；就方法论气质来说，它反而离自己的批判对象KKV更近。TTC虽然就自己与KKV的差异对KKV展开了批判，但也同时捍卫了它和KKV所共享的一些理念。而从社会学方法论的总体视野来讲，TTC和KKV所共享的那些理念，可能是更为重要

的，也是更成问题的。

TTC把量化研究与质性研究描述成“两种文化”未尝不可，但是从社会科学哲学的角度看，TTC和KKV恰恰表现了同一种文化，即逻辑实证主义影响下的美国社会科学。这也正是TTC在政治学和管理学领域的影响力和传播力远大于社会学界的原因，因为政治学和管理学没有社会学那么深厚的欧陆思想的方法论积淀（从负面来说，也可以叫包袱）。从TTC一书举为例证的一些经验研究中，我们就不难看出上述不足。许多使用QCA方法的政治学家所选择的研究题目与量化研究很相似，比如民主和经济增长的关系。单从选题来说，这种“价值介入”并不成问题。然而我们能够看到，按照QCA方法开展此类研究的质性研究者往往陷入了与主流量化研究一样的陷阱，即在讨论截面的“客观证据”时脱离了宏观的历史背景，从而表现出严重的“物化”思维。同KKV一样，TTC对于理论与经验之间的关系理解得过于简单，经验研究往往最终成为了某些“中层理论”的注脚，以致很难对社会科学的事业做出有重大意义的理论贡献。

五、结语

方法论之争是当今中国社会学界的重要现象，近年来社会学界的许多针锋相对的讨论都发生在方法论领域，如谢宇（2018）、贺雪峰（2020）、渠敬东（2016）、翟学伟（2018，2020）、梁玉成（2018）、周晓虹（2020）、郑莉（2019）、王苍龙（2021）等人参与的本土化议题，潘绥铭（2016a，2016b）与刘林平（2016）参与的大数据议题，渠敬东（2019）、耿曙（2019）、张静（2018）等人参

与的个案方法议题，赵鼎新（2015，2016，2020a，2020b）、李连江、叶启政（2020）等人参与的概念与机制议题。考虑到目前大陆学术界本身缺少批判精神的文化背景，上述有些火药味的讨论透露出了非常强的对抗情绪。

量化研究与质性研究的对抗，正处于上述对抗的核心位置。然而当我们将“从KKV到TTC”这一段美国学界的方法论历史与当前中国社会学界的方法论对抗并置在一起时，就会感到有些错位。其原因显然，目前在大陆社会学界倡导质性研究的学者鲜有QCA的背景。由此可见，尽管人们常常就美国的社会科学传统展开论辩，但无论是倡导还是抵触，都表现得有些片面，或至少存在着一些对于“他者”的认识偏差。这是一个有趣的知识社会学现象。

总的来看，从KKV到TTC的这段论辩对当今中国的方法论研究者可以有若干启发：第一，在批判量化研究时，质性研究者需要学会清晰表达自己的方法论立场，而不是以含混的概念来进行无效对话；第二，量化研究者需要认识到，量化研究原本在方法论上的“清晰”优势已经受到了实际的撼动，只不过由于学术传播的滞后性和片面性，国内的量化研究者在方法论上得以回避了实质挑战；第三，要想在方法论讨论中实现实质性的沟通与理解殊为困难，方法论争论有着极强的“文化性”，保持开放态度或许是学者的最佳选择；第四，社会学方法论这一领域自诞生之日起就被赋予了理论性和工具性两重特质，也因此陷入了理论性与工具性的矛盾中。不同的方法论流派在应对这一矛盾时所选择的路径有很大差别，有的无视这一矛盾，有的则陷入矛盾中不能自拔。方法论研究者一方面需要保持理论的反思性，另一方面也要致力于对工具性的追索，寻找一种持衡的知识姿态。

可以想象到，随着进一步的学术交流和传播，中国社会学界势必会有越来越多的研究者受到TTC一书的影响，运用QCA方法开展质性研究的学者也会越来越多。但是，在认识和传播TTC一书的方法论价值时，我们需要摆脱常见的"翻转"思维，即单纯将TTC视为KKV的方法论对立面。这种思维依然潜藏着研究方法领域中一直以来的"拿来主义"倾向，"从KKV到TTC，无非是学习一种新工具"，这种认识才是当前中国社会学研究方法的真正危机所在。近年来"方法主义"的盛行给中国社会科学带来了虚假繁荣，我们需要破除对于方法的"工具化"思维，才有助于建立中国社会科学家的主体性。TTC与KKV在方法论立场上的分歧与共识，为我们开展上述反思提供了有益起点。

参考文献

耿曙，2019，《从实证视角理解个案研究：三阶段考察渠文的方法创新》，《社会》第1期。

何俊志，2013，《比较政治分析中的模糊集方法》，《社会科学》第5期。

贺雪峰，2020，《本土化与主体性：中国社会科学研究的方向——兼与谢宇教授商榷》，《探索与争鸣》第1期。

加里·格尔茨、詹姆斯·马奥尼，2016，《两种传承：社会科学中的定性与定量研究》，刘军译，上海：上海人民出版社。

加里·金、罗伯特·基欧汉、悉尼·维巴，2014，《社会科学中的研究设计》，陈硕译，上海：格致出版社。

李蔚、何海兵，2015，《定性比较分析方法的研究逻辑及其应用》，《上海行政学院学报》第1期。

梁玉成，2018，《走出<走出中国社会学本土化讨论的误区>的误区》，《新视野》第4期。

刘林平、唐斌斌、蒋和超，2016，《大数据有“原罪”吗?——与潘绥铭教授商榷》，《社会科学文摘》第8期。

马克斯、贝努瓦·里候科斯、查尔斯·拉金，2015，《社会科学研究中的定性比较分析法——近25年的发展及应用评估》，臧雷振译，《国外社会科学》第6期。

潘绥铭，2016a，《生活是如何被篡改为数据的?——大数据套用到研究人类的“原罪”》，《新视野》第3期。

——2016b，《再论生活是如何被篡改为数据的——回应刘林平教授的质疑》，《新视野》第4期。

渠敬东，2016，《破除“方法主义”迷信：中国学术自立的出路》，《文化纵横》第2期。

渠敬东，2019，《迈向社会全体的个案研究》，《社会》第1期。

释启鹏，2016，《时间中的定性比较分析：TQCA与TSQCA的发展》，《比较政治学研究》第1期。

唐世平，2015，《超越定性与定量之争》，《公共行政评论》第4期。

王苍龙，2021，《中国社会学本土化的两个立场与两条路径——兼与谢宇、贺雪峰教授商榷》，《社会科学》第1期。

王水雄，2001，《亲和性机制或"虚假"命题——新教伦理与资本主义精神的逻辑缺憾》，《社会学研究》第2期。

夏鑫、何建民、刘嘉毅，2014，《定性比较分析的研究逻辑——兼论其对经济管理学研究的启示》，《财经研究》第10期。

谢宇，2018，《走出中国社会学本土化讨论的误区》，《社会学研究》第2期。

叶启政，2020，《从因果到机制》，台北：群学出版社。

翟学伟，2018，《社会学本土化是个伪问题吗》，《探索与争鸣》第9期。

——2020，《为什么社会学本土化要“窄化”? ——对周晓虹教授的几点回应》，《开放时代》第5期。

张静，2018，《案例分析的目标：从故事到知识》，《中国社会科学》第8期。

赵鼎新，2015，《社会科学的研究困境：从与自然科学的区别谈起》，《社会学评论》第4期。

——2016，《论“依法抗争”概念的误区：对李连江教授的回应》，《社会学评论》第1期。

——2020a，《论机制解释在社会学中的地位及其局限》，《社会学研究》

第2期。

——2020b,《机制解释的诠释学挑战和回应》,《社会学评论》第6期。

郑莉,2019,《中国社会学研究"规范化"与"本土化"之知识论探析》,《天津社会科学》第4期。

周晓虹,2020,《社会学本土化:狭义或广义,伪问题或真现实——兼与谢宇和翟学伟两位教授商榷》,《社会学研究》第1期。

Brady, Henry E. & David Collier 2010, *Rethinking Social Inquiry Diverse Tools, Shared Standards.* Second edition. New York, NY: Rowman & Littlefield Publishers.

Collier, David, Jason Seawright & Gerardo L. Munck 2010, "The Quest for Standards: King, Keohane, and Verba's Designing Social Inquiry." In Henry E. Brady & David Collier (ed.), *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* Second edition. New York, NY: Rowman & Littlefield Publishers.

Goertz, Gary & James Mahoney 2012, *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences.* Princeton, NJ: Princeton University Press.

Johnson James 2006, "Consequences of Positivism." *Comparative Political Studies* 39(2).

King, Gary, Robert O. Keohane & Sidney Verba 1994, *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

McKeown, Timothy J. 1999, "Case Studies and the Statistical Worldview: Review of King, Keohane, and Verba's Designing Social Inquiry: Scientific Inference in Qualitative Research." *International Organization* 53(1).

Ragin, Charles 1987, *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies.* Berkeley, CA: University of California Press.

Ragin, Charles C., S. E. Mayer & K. A. Drass 1984, "Assessing Discrimination: A Boolean Approach." *American Sociological Review* 49(2).

Stevens, Stanley Smith 1946, "On the Theory of Scales of Measurement." *Science* 103(2684).

吕鹏,李蒙迪,张卓．基于沙堆模型的国家动力学研究——以中国历史为例(前221年到1912年)[M/OL]//赵联飞,赵锋．社会研究方法评论:第2卷．重庆:重庆大学出版社．

基于沙堆模型的国家动力学研究[1]

——以中国历史为例（前221年到1912年）

吕 鹏 李蒙迪 张 卓[2]

摘要:关于帝国兴衰，长期存在的“超稳定结构”理论并不能呈现动态的历史过程。复杂性学科的“自组织临界性”理论，能够模拟帝国兴衰的动态过程，并对“超稳定结构说”形成补充。自然系统（沙堆模型）与社会系统类似，都在自动、自发地趋近临界状态，我们可以通过社会计算模型揭示帝国兴衰的动力学机制。研究目标，是精准拟合中国大历史，即秦汉到明清的2132年的农业帝国兴衰演变史，包括朝代数量、朝代寿命等客观指标，使用具体智能体ABM模型，建构代表帝国社会有机体的沙堆模型。通过参数遍历仿真，求解最优的参数解。基于最优解，重复模拟1000次，求取均值获得稳健性结果。研究表明：第一，沙堆模型具有拟合有效性。中国历史上有22个朝代，沙堆模型可以“涌现”相同数量

---

①本文部分内容曾以英文发表。相关信息如下：Peng Lu，Hou Yang，Mengdi Li，Zhuo Zhang. The Sandpile Model and Empire Dynamics. *Chaos*，*Solitons and Fractals*，2021，143，110615。

②作者简介：吕鹏，中南大学教授、社会计算研究中心主任、青年长江学者，联系方式：lvpengccps@vip.qq.com；李蒙迪，中南大学社会学系博士；张卓，中南大学社会学系博士，联系方式：210201011@csu.edu.cn。

的朝代或帝国，而且数量遵循正态分布（平均值是22）。第二，沙堆模型具有重复稳健性。朝代寿命遵循幂律分布，模拟结果与真实历史呈现高度吻合。第三，沙堆模型具有可推广性。同样可以刻画非22帝国情况。22朝代是公认标准，但历史存在非线性、偶然性。因此需要拟合数量不等于22的情况。舍弃历史上1~3个短寿朝代，构建19、20、21组历史与仿真朝代队列，依然实现高精确度拟合。在23、24、25个仿真朝代中，舍弃1~3个短寿者，构成22组历史与仿真朝代，拟合度同样良好。沙堆模型对帝国兴衰历史、当前中美大国博弈，具有借鉴与指导意义。

关键词：智能体建模（ABM）；帝国动力学；历史周期；组织临界性

Abstract: According to the self-organized criticality theory, the sandpile model is built to investigate the evolutionary dynamics of empires in the history of China. The methods of agent-based modeling and simulations are applied to capture empires' mechanism of rising and falling cycles, and to obtain the observed life cycle pattern of empires in history. Under the self-organized criticality theory, natural systems and human empires systems have similar structures and mechanisms, which makes systems reach the critical states automatically. Therefore, the rising and falling dynamics of empires can be reflected by the sandpile model as well. With the sandpile modeling and simulations, the optimal solution of parameters can be found, based on which the satisfactory fitness of results can be achieved. Under the optimal solution, we run the simulations for 1000 times to check the fitness and robustness. First, the number of empires can be matched. There were 22 empires in the history of

China, and the same number of empires can be obtained via sandpile model simulations, and the amount of empires follows the normal distribution with the mean of 22 empires. Second, the distribution of empire durations follows the power-law distribution, for both simulated and historical empires. Third, for less than 22 simulated empires, we drop empires with tiny durations in history to compare the 19, 20, and 21 pairs of counterparts respectively, and the fitness can be guaranteed as well. Finally, for more than 22 empires, we drop simulated empires with tiny durations to compare the 23, 24, and 25 pairs, the matching degree is satisfactory as well. It indicates that multiple simulations have more robust and stable outcomes than one, even the best, simulation.

Key words: Sandpile Model; Agent-Based Modeling (ABM); Empire Dynamics; Rise and Fall Cycles; Self-Organized Criticality

一、引 言

帝国被定义为一种主权政治实体（Barkawi & Laffey，2002）。尽管历史上的帝国有不同的持续时间和成长-衰落过程，但“生命周期”模式是稳健的（Ferguson，2010）。人类历史上，帝国生命周期多包含爆发、成长、巅峰、消失等典型的阶段。学术界的帝国发展动力学研究，分为质性社会研究和跨学科的定量研究。如下所示。（1）质性社会研究。研究重点是对帝国某一具体方面进行探讨。通过概念模型研究帝国的发展动力学，而非借助模型化方法来解析帝国的动力学。例如，威默分析了西方1816—2001年期间的所有战争，认为帝国的解体、民族国家的形成是区域冲突的主要

原因（Wimmer & Min，2006）。诺尔库斯使用帝国的概念来分析中世纪的立陶宛公国（Grand Duchy of Lithuania，GDL），并根据军事力量、领土结构、霸权范围和长途贸易的理论框架，将立陶宛公国定义为一个帝国实体（Norkus，2008）。奎罗加以自由人（区别于奴隶概念）为分析单位，探讨意大利帝国的兴衰，并指出自由人在商业、手工业或银行构建的社会网络的重要性（Quiroga，2010）。斯泰因梅茨提出了一个关于帝国规模的理论，并区分了社会场、子场和社会空间等概念（Steinmetz，2016）。（2）定量研究方法。区别于传统的社会历史研究，多出现在跨学科研究。在数学模型方面，塔塞佩拉提出了一个简单的曲线模型来模拟帝国的增长动态（Taagepera，1978）。该模型包括一个缓慢的开始、快速的扩张、稳定的最大规模三阶段（Taagepera，1979，1978）。也有学者通过分析人口动态模型，建立数学模型来研究人类社会或帝国演化（Cheon & Poghosyan，2017）。研究发现，螺旋轨道和振荡可以解释帝国的兴衰演变，社会进化的文化土壤也可以被计算建模。蔡斯-邓恩等人探讨了青铜时代以来的兴衰因果模式，认为人口、领土规模可以解释帝国动力学（Chase-Dunn et al.，2006）。在皇帝在位时间方面，赫马拉泽认为，罗马、中国皇帝的统治时间都服从指数分布（Khmaladze et al.，2007）。随后，蔡斯-邓恩等人利用包括人口、领土规模、央地关系等因素，来研究帝国的动态发展（Chase-Dunn et al.，2012）。阿尔贝斯曼研究了全球帝国寿命的分布特征，发现了“无记忆指数分布”的规律（Arbesman，2011），该分布与物种、公司的分布规律类似。马尔凯蒂和奥苏贝尔发现帝国崩溃的动力学模式呈现S型（Logistic 曲线），此外，物流运输速度决定了帝国疆域的大小（Marchetti & Ausubel，2012）。

弗格森采用了文明或帝国的生命循环视角，认为当前的全球大国与历史上的全球帝国有类似的发展轨迹（Ferguson，2006）。叶戈尔夫建构了一个动态优化模型，包括地理、经济、边界动态、运输成本、目标和帝国的演变等多个维度。它表明，多重性均衡可以解释在历史上大国和小国共存的原因（Yegorov et al.，2019）。

梳理已有的模型研究，尚未发现“自组织临界性”的仿真模拟应用，尤其是在用沙堆模型研究帝国动力学方面。在帝国动力学领域，传统的社会研究方法一直占据重要位置，通常采用定性研究方法。其研究往往具有局部的精深性，但不是宏观、系统、整体层面的有机体研究。跨学科的量化研究方法，采用数学模型与定量模型，能够解释过程与规律。但属于静态性研究，缺乏动态性，不能还原帝国动力学的完整过程。我们的研究模式，属于第三种方法，即计算社会科学方法。我们以中国历史为例，构建基于多智能体的计算仿真模型，探索帝国生命周期的动力学机制。研究者目前多是从地理因素（Chase-Dunn，1990）、经济发展（Gills & Frank，1992）、人口增长（Turchin，2005）、军事战争（Eckhardt，1990）、农业生产（Shimada & Koyama，2015）等层面，研究帝国兴衰的动力机制（Hall & Chase-Dunn，1993；Sinopoli，1995）。基于“长周期理论”，在全世界范围内，帝国在“历史-政治”交互作用下，均呈现周期性更迭现象（Rosecrance，1987）。因此，中国历史上的帝国周期律，也呈现出类似的分布规律。鉴于在生命周期方面，中国与全球的帝国存在相同的模式，那么生命周期率背后必然存在着稳固的动力学机制。我们虽然有了一些理论模型、数学模型。通过这些模型，能够在一定程度上理解帝国的形成（崛起）和崩溃（衰落）。但其都过于理论化，离“高精度复现历史动态过程”的目标

相距甚远。因此，研究帝国生命周期需要更广泛的跨学科研究视角。在此方面，计算机仿真方法，具有显著的研究优势。多智能体模型（Agent-Based Modeling）是目前最前沿的计算社会科学工具，它可以精确复现、精准还原、动态反算（backcalculate）帝国的生命周期过程，并进行可视化。这种方法还没有被广泛用于探索帝国的兴衰机制。目前，尚未发现有从计算模型角度探讨生命周期内在机制的研究。因此，我们从“自组织临界性”出发，基于沙堆模型建立智能体模型。探索历史上帝国兴衰演化的科学内核。

二、自组织临界性理论（Self-Organized Criticality）

巴克等人最早提出自组织临界性（Self-Organized Criticality），这种性质存在于诸多自然、社会系统中（Bak et al.，1987）。根据自组织临界理论，一个动态系统会自发地朝临界状态演化。这是一个自动化的过程，而且具有稳健性，即这种临界状态不受系统中的复杂性、非线性的干扰。当偏离临界状态时，系统又会自动返回临界状态（Bak et al.，1988）。自组织临界性被认为是解释复杂系统的一般化机制的关键概念。在数学上，复杂系统行为遵循简单的模式。自组织临界性是耗散结构系统的核心行为演化模式，有两个基本特征：时间效应中的1/*f*噪声，以及空间结构演化中的尺度不变的自相似性（Teran，2001）。沙堆模型是自组织临界性的典型研究工具。3年之后，海尔德等人（Held et al.，1990）进行了一系列沙堆实验。选择1~1.25毫米直径的沙粒，在半径为2厘米的圆盘上建造沙堆。当倾斜角接近临界状态，沙堆停止增长，加入的颗粒数约等于落在外面的颗粒数。接近临界状态的时候，新加入的沙粒所带来的影响

是不可预测的。沙粒可能会留在沙堆上，引起小的滑动，也会导致大规模雪崩。然而，学者们进一步发现，雪崩的规模与频率的关系遵循幂律分布（Held et al.，1990）。在沙堆雪崩过程中，塌陷的沙粒数是雪崩规模（S），持续时间被称为崩塌时间（T）。雪崩发生的概率（D）、雪崩规模（S）与崩塌时间（T）的关系满足幂律分布。

在多种因素的相互作用下，系统会自动地达到临界状态，不需要进行人为的参数调整。这意味着，在该系统中，即使是一个小事件也会引发连锁反应，导致大规模级联崩塌。虽然系统中的小事件多于大事件，但各种尺度的连锁反应是动态系统的重要组成部分。因此，幂律分布是证明系统的自组织临界性的统计方法（Bak & Chen，1991）。自组织临界性理论已用于解释地震学（Bak & Tang，1989）、地质学（Diodati et al.，1991）、湍流（Sapozhnikov & Foufoula-Georgiou，1991）、天体物理学（Garrido et al.，1996）、计算机网络（Yuan et al.，2000）、交通（Elmer，1997）、生物学（Babcock & Westervelt，1990）、森林火灾（Malamud et al.，1998）、经济学（Stauffer & Sornette，1999）中的复杂性现象。例如，人们发现，许多元胞自动机模型都呈现类似的自组织临界性，如地震、生命游戏、森林火灾、股市震荡等。例如，Bak 等人发现，地震时释放的能量分布也遵循幂律分布（Bak & Tang，1989）。地壳处于运动状态，发生缓慢变形，导致岩石中的应力积累。如果岩石在某处不能承受能量，能量就会释放出来。如果系统环境能够承受这种能量的增加，它将是一个孤立的事件。但如果此处也恰好处于临界状态，就会发生连锁反应。孤立、零星事件将被连接成宏观涌现。

“自组织临界性”理论已扩展到社会和历史研究方面，这为帝国动力学研究做了很好的铺垫和注脚。罗伯茨和特科特（Roberts

& Turcotte，1998）研究历史上的战争，认为世界秩序也具有类似的性质（Elmer，1997）。世界上有一些局部的小冲突，这些冲突有的可能会发展成大的战争，但大多数的影响力仅限于局部。世界秩序也是一个自组织系统，这是一种强大的客观规律，与各国的控制努力无关（Roberts & Turcotte，1998）。皮科利等发现，人类冲突中的暴力事件可能与阈值机制有关，可以用自组织临界性来描述（Picoli et al.，2014）。岛田和小山在社会动力学系统中发现了危机征兆，该征兆表明系统已经准备好进行大规模的转变。所以检测自组织的临界性可以用来诊断社会的转型潜力（Shimada & Koyama，2015）。朱可夫等人分析了俄罗斯过去两个世纪的人口和粮食价格，也验证了1/*f*噪声规则（Zhukov et al.，2016）。可见，当我们研究人类社会复杂系统（如帝国）的兴衰时，自组织临界性理论和模型将提供一个有力的分析角度。

三、构建理论类比

沙堆和帝国都具有自组织临界性。它们有相似的层次结构和反应过程。因此，人类冲突模式可以和自然现象中的自组织临界性进行类比。具体如下所示。

（一）相似的社会结构分层

自中国、埃及和印度等人类文明建立以来，社会等级制度或社会分层现象一直存在于人类社会（Yao，2005；Li，1996）。我国农业社会，大致分为“士农工商”四大等级与类别。如清朝时，社会被划分为社会官僚、乡绅、工人阶级、下层阶级（Zhao-hui，

2007）。类似地，等级制度在古埃及、希腊和印度也出现过（Triandafyllidou & Veikou，2002；Basu，2012）。在古代印度，种姓制度（婆罗门、刹帝利、吠舍和首陀罗）划分了不同的社会等级，这种制度一直延续到现在。在古埃及，除了法老（国王）之外，人们被分为宰相、贵族、祭司、文士和士兵等（Mahaffy，1899）。沙堆结构类似于帝国的社会结构，底部的沙粒象征着被统治的工匠与农民。他们为整个社会生产能量（财富）。这是沙堆的"基石"，他们支撑着整个沙堆系统。中间的沙粒，代表着商人和士兵，上面指统治阶层（文官集团、国王等）。具有临界状态的沙堆，呈现金字塔形状。添加沙粒的过程，类似于帝国社会的发展。社会分层和社会矛盾会逐渐积累，直至爆发系统性动荡。

（二）相似的"矛盾积累—总爆发"过程

沙堆的积累和崩塌过程与帝国的崛起和衰落过程相似。如，西诺波利（Sinopoli，1994）将帝国成长到衰落的过程概括为三个阶段，即扩张、巩固和崩溃（Sinopoli，1994）。可见，帝国的最终命运是走向死亡、崩溃。但是，没有人知道何时发生。帝国的系统性崩溃，具有结果的确定性、过程的不可预测性的双重性质。用沙堆模型来做研究，其整个过程也分为三个阶段：开始—输入—崩溃。模型初始化，我们设定世界（系统）范围内初始沙粒的数量及其分布。然后，在单位时间内，我们将沙子一粒一粒地添加到沙堆中。首先，在初始状态下，沙子会留在它们所在的位置（Patch）附近。然后，它们被一粒一粒地堆积起来，形成一个缓慢倾斜的沙堆系统。最后，当坡度变得过于陡峭时，沙子会滑下来，导致局部雪崩（坍塌）的发生。随着更多沙子的加入，坡度变得更陡，形成更多的大雪崩或

小雪崩。下一时刻的雪崩或坍塌，具有不可预测性。这个过程持续进行、不以人的意志为转移。沙堆系统是一个具有自组织行为的演化系统，其演化动力学是由内因（沙堆结构）决定的。尽管导致王朝兴衰的因素是复杂多样的，包括经济活动、文化制度、技术水平、地理环境、军事力量等（Förster，1988），但是王朝的周期性演变，也主要是由内生的因素和驱动力（动态演化的社会结构）决定的。

（三）二者都具有自组织临界属性

基于结构、过程相似性，帝国社会和沙堆模型也具有相同的自组织临界性的系统性质。在沙堆中，自组织过程体现为随机的累加、不可预测的崩塌，这是对帝国的成长、稳定和崩溃的恰当隐喻。帝国的运行，也是动态的加压，最终走向系统性崩溃。当加入的沙子数量与外面的沙子数量相平衡时，沙堆停止增长，系统就达到了临界状态。当更多的沙子被添加到已经处于临界状态的沙堆中时，会引起不同规模的雪崩现象。在临界状态时，沙堆保持一个恒定的高度和坡度。如果坡度小于亚临界状态的临界值，雪崩将小于临界状态下产生的雪崩。如果斜率大于临界值，雪崩的规模将大得多。超临界堆的崩溃，将在其达到临界状态之前发生。亚临界和超临界的沙堆自然会被发展到临界状态。帝国的发展可以被认为是系统的进化。当达到临界点时，帝国就会发生“相变”。系统的自组织临界性已被应用于分析帝国、文明、政治或群体行为。如，伊博勒使用该视角研究人类的定居文明何以出现（Iberall，1987）。葡托盖里分析了城市和其他大型文化实体的周期和系统（Portugali，1997）。因此，相关的自然科学与社会科学研究，证明了社会组织、帝国社会呈现的自组织临界性，与自然系统的相关特征具有一致性（Iberall，1984）。

四、研究方法与模型建构

我们以中国历史为例，试图描绘、反算、预测人类社会的历史进程。如附表1所示，在中国大历史的2132年进程中，一共出现22个帝国或朝代。本研究中，沙堆模型的建构是由Netlogo软件完成。核心思路是：设置全社会的危机阈值，实时收集沙堆坍塌比例，生成或识别帝国朝代，寻求拟合历史朝代的最优仿真参数解。

（一）多智能体行为规则设置

Netlogo是一个用于模拟自然和社会现象的编程环境。有动态智能体（turtles）和静态智能体（patches），都属于智能体类别（agents）。静态智能体，用于模拟不可移动物体，如建筑物、山脉、河流、土地和植物。动态智能体，可以模拟个人、动物、粒子、汽车等。鉴于我们的核心目标，是匹配前221年至1912年的帝国兴衰历史。我们在模拟中设定2132次迭代，每次tick代表1年。在沙堆模型中，我们用静态的Patch设定帝国疆域，动态的主体设定为沙粒。帝国被设定为一个正方形，总面积为$S = 41 \times 41 = 1681$。在每个时间t，都会有一颗沙粒加入模型中。每个Patch上的沙粒初始数量为1—4颗。每个时间t内，一颗沙粒会被随机添加到一个Patch上。当沙粒的数量达到某Patch的临界阈值时，就会发生塌陷。4颗沙粒将从此Patch平均地转移到4个邻近的Patch。当前Patch的沙粒变成1，开始重新累积沙粒。然后，邻近的Patch可能变得过载，导致新一轮坍塌。临界阈值由界面的滑动条“Threshold”控制，模型设置为5。整个世界的边界不相连的，一旦沙粒离开边界，就会消

失。在每个tick，我们将一开始的沙堆崩塌到完全不再有新崩塌的整个过程，称为一次雪崩或坍塌。塌陷Patch的数量*Ct*被视为雪崩的绝对量大小，即涉及的沙粒总数。崩塌的百分比，被计算为Pt=Ct/S。S是Patch总数。在每个时间t，我们计算并记录P_t，与帝国的坍塌阈值进行比较。

（二）帝国演化过程与朝代信息提取

沙堆模型刻画了人类帝国的动态演变过程。模型初始化，代表一个王朝开始了。对于沙粒的初始设置，每个Patch上有4颗沙粒。所有Patch的崩塌阈值是5，该阈值在全过程保持不变。这意味着，所有Patch或区域都有相同的临界性。因此新放入的沙粒可以迅速产生崩塌。沙堆（帝国）的动态过程如下：在每个时刻，一颗沙粒被添加到某一个Patch上，位置是完全随机的。当Patch上的沙粒数量达到临界阈值5时，就会产生直接、间接崩塌。就四个Patch而言，每个都从中心Patch的坍塌中获得1颗沙粒，这将或然地引起进一步的级联崩塌。受崩塌影响的Patch在视觉上会标记成银白色，便于统计系统宏观状态。Netlogo中，每个智能体都有一个唯一的ID，表示为i。这是一个动态的系统过程，我们用公式（1）来反映Patch的变化。对于一个任意Patch，沙粒的数量在时间t表示为z_{it}。在下一个时间$t+1$，其上的沙粒数为$z_{i,t+1}$有三种可能性。如果是中心Patch，沙粒的数量在时间上t满足阈值5时，将失去、转移4颗沙粒到4个邻居Patch；若是已满足阈值的中心Patch的邻居，或是其沙粒数未满足阈值5，本轮沙粒数增加1；否则，数量不变。

$$z_{i,t+1}=\begin{cases} z_{it}\text{ifoutside} \\ z_{it}+1\text{ifneighbor} \\ z_{it}-4\text{ifcenter} \end{cases} \quad (1)$$

在每个时间t，我们计算崩塌的百分比（$P_t = C_t/S$）。中国、世界历史上，该百分比P_t反映了社会不稳定和系统混乱的程度。这种社会不稳定的系统性混乱造成了帝国的崩溃。对于强大而稳定的中央政权，较弱的混乱能够被控制或处理。然而，更严重的混乱和不稳定，将必然会推翻现有的政权或帝国。因为这种混乱超过帝国的控制力范围。我们因此设置阈值机制（百分比）刻画帝国的控制能力。在每个时间点，我们计算出P_t，并与帝国的控制危机的能力阈值P_T比较。如果是$P_t \geqslant P_T$，这意味着当前朝代将要灭亡，新的王朝成立。如果$P_t < P_T$，则混乱和风险仍在帝国的控制能力之中，王朝继续存活。根据此规则，我们可以在整个2132年（Ticks）动态演化的时间序列中切割（划分）出若干王朝。如附表1所示，我们通常认为有22个朝代（帝国），因此我们要从2132年模拟中提取22个左右的帝国。

（三）寻找最佳参数和解决方案

核心的任务，是求解最佳的参数组合即$Paramters^*(\cdot)$，产生最符合历史的仿真结果。在数学上，22个朝代可用$(Y_1,Y_2,Y_3,\cdots Y_i,\cdots Y_{21},Y_{22})$表示，是一个有序的帝国寿命序列，总和为2132年。持续时间（国祚）是帝国的关键信息。我们将历史上的国祚数量集合定义为真实目标函数即$f_{real}(\cdot)$。在模型中，不同处理危机的能力阈值（P_T），仿真出来的帝国寿命、数量也不同。并非每一次仿真都能产生良好的拟合效果，有很大比例的仿真是不理想的。对每一次仿真，我们重复100次以获取稳健的均值结果，如$(\widehat{Y_1},\widehat{Y_2},\widehat{Y_3},\cdots \widehat{Y_k},\cdots \widehat{Y_{21}},\widehat{Y_{22}})$。对应地，模拟得到的帝国寿命结果的集合，构成了仿真结果函数$f_{sim}(\cdot)$。为计算拟合优度，我们对真实帝国寿命$(Y_1,Y_2,Y_3,\cdots Y_i,\cdots Y_{21},Y_{22})$和模拟帝国寿命$(\widehat{Y_1},\widehat{Y_2},\widehat{Y_3},\cdots \widehat{Y_k},\cdots \widehat{Y_{21}},\widehat{Y_{22}})$进行排序，然后组成模拟–历史帝国配

对，计算匹配度。对不同的参数设置进行模拟，模拟结果的拟合程度不同。在中国，大多数朝代都有明确的划分（清晰的历史传承与记录），但也有一些混乱的时期，我们将混乱期视为一个朝代。例如，五代十国（907—960）有15个并存、割据的政权，三国有3个政权。此外，国祚匹配的误差在一定程度内可接受（小于10年），不能对帝国整体产生实质性影响。

我们仿真遍历所有可能的参数组合，持续计算拟合度 $\Delta = f_{sim}(\cdot) - f_{real}(\cdot)$。目标是获得最优参数解，使得历史与模拟之间的匹配度最高（误差最小）。如何测量匹配度？我们考虑三个标准：（1）数量匹配是基本标准。模拟帝国的数量应该与真实历史相匹配，即 $max(i) = max(k) = 22$。模拟的期望值（平均值）应与22相匹配，即 $E(k) = max(i) = 22$。如果在最优解下进行多次模拟，那么观察数 k（模拟获得的帝国数量）应遵循正态分布，均值应为22，以100%地匹配历史记录；（2）进阶标准是分布匹配。考虑到要匹配22个帝国，其寿命的分布也应与历史相匹配，包括离散分布、概率密度分布。（3）最高标准是模拟和历史朝代的国祚之间实现“一对一”高精度匹配。最佳目标函数 $Paramters^{*}(\cdot)$ 使用公式（2）求解。一旦确定了最优参数解，还将追加1000次模拟，以验证其稳健性。

$$Paramters^{*}(\cdot) = Argmin[f_{sim}(\cdot) - f_{real}(\cdot)] = Argmin\sum_{1}^{22}\frac{\left(Y_i - \widehat{Y_i}\right)^2}{n-1} \quad (2)$$

$$s.t.\sum Y_i = 2132$$

五、研究发现与结果呈现

经过仿真模拟与结果考察，确定了最优参数组合（最优解）。

帝国克服危机的能力临界阈值等于0.563，初始沙子数量为4，坍塌阈值等于5。最优解的存在证明，自组织临界性可用于研究自然科学、社会科学、历史科学的跨系统现象。基于该理论的沙堆模型，能够揭示人类历史上帝国兴衰的动态演化机制。

（一）模拟结果的分布拟合度

在最佳参数*Paramters**(·)下，进行1000次模拟。图1展示了沙堆模型在最佳参数下的模拟结果。由于历史进程和沙堆模拟的过程具有随机性、动态性和复杂性。除了符合历史史实的22个帝国之外，我们不可避免地得到12~33不等的模拟结果。但是整体来看，平均值

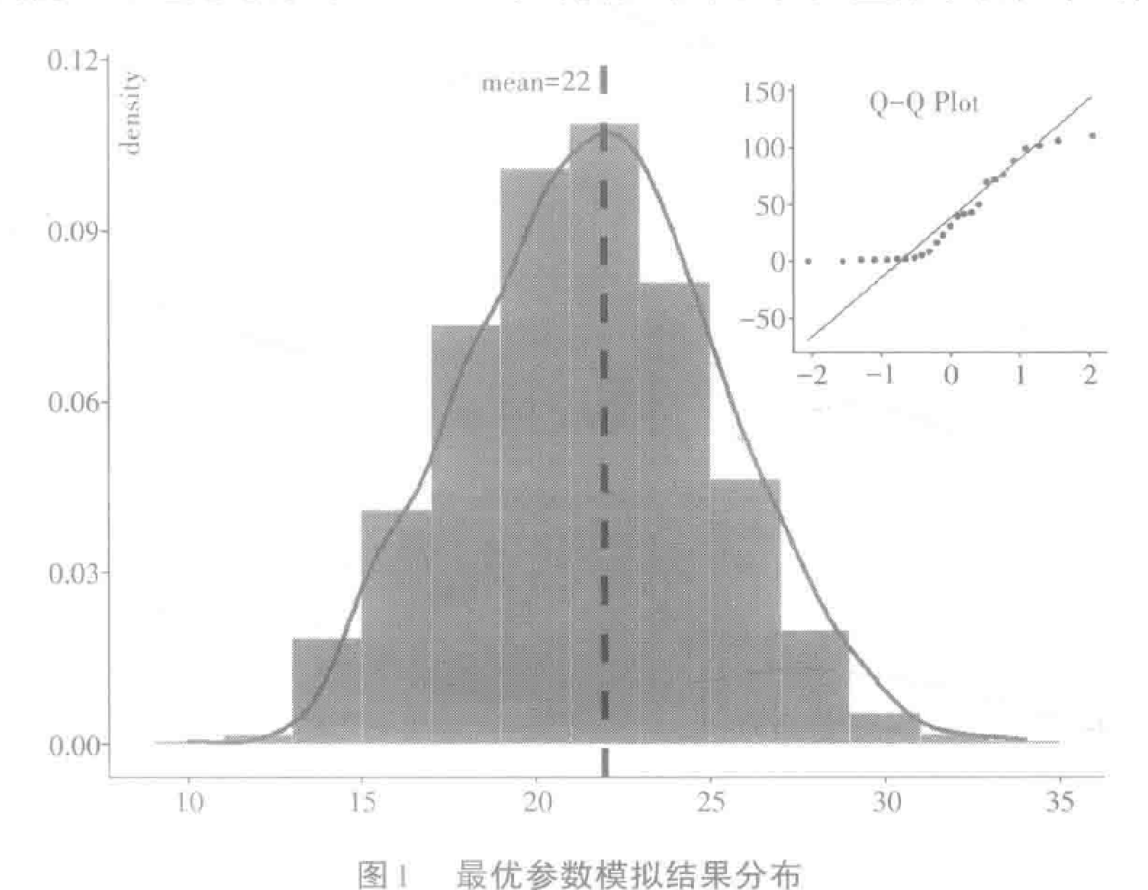


图1　最优参数模拟结果分布

注：主图显示了沙堆模型在最优参数下1000次模拟的帝国数量的概率密度分布。*x*轴坐标表示模拟帝国的数量，*y*轴坐标表示概率密度每次模拟生成的帝国数量不同，我们有1000个观察值。模拟的朝代数量从12到33不等，但是平均值为22（SD=3.598），吻合历史。子图Q–Q正态图检验正态性。

是22（SD=3.598），这证明了沙堆模型的有效性，以及最优解的稳健性。图1中，我们将观测到的不同朝代结果可视化。其中，N=15（3.10%），N=16（4.00%），N=17（4.20%），N=18（7.70%），N=19（7.00%），N=20（10.00%），N=21（10.20%），N=22（11.10%），N=23（10.70%），N=24（8.90%），N=25（7.30%），N=26（5.00%），N=27（4.30%），N=28（2.30%），N=29（1.70%）。概率密度函数的均值为22，对称性、趋中性良好。Q-Q正态图进一步表明模拟的数量分布接近正态分布。因此，图1表明沙堆模型能够模拟和反算历史上帝国的兴衰动态。

图2考察了寿命分布。无论是真实历史、最佳一次模拟、平均模拟结果，均呈现右长尾分布。塔尔戈夫斯基（Targowski，2009）认为，历史上的人类文明与帝国寿命有规律性，大致符合某种指数

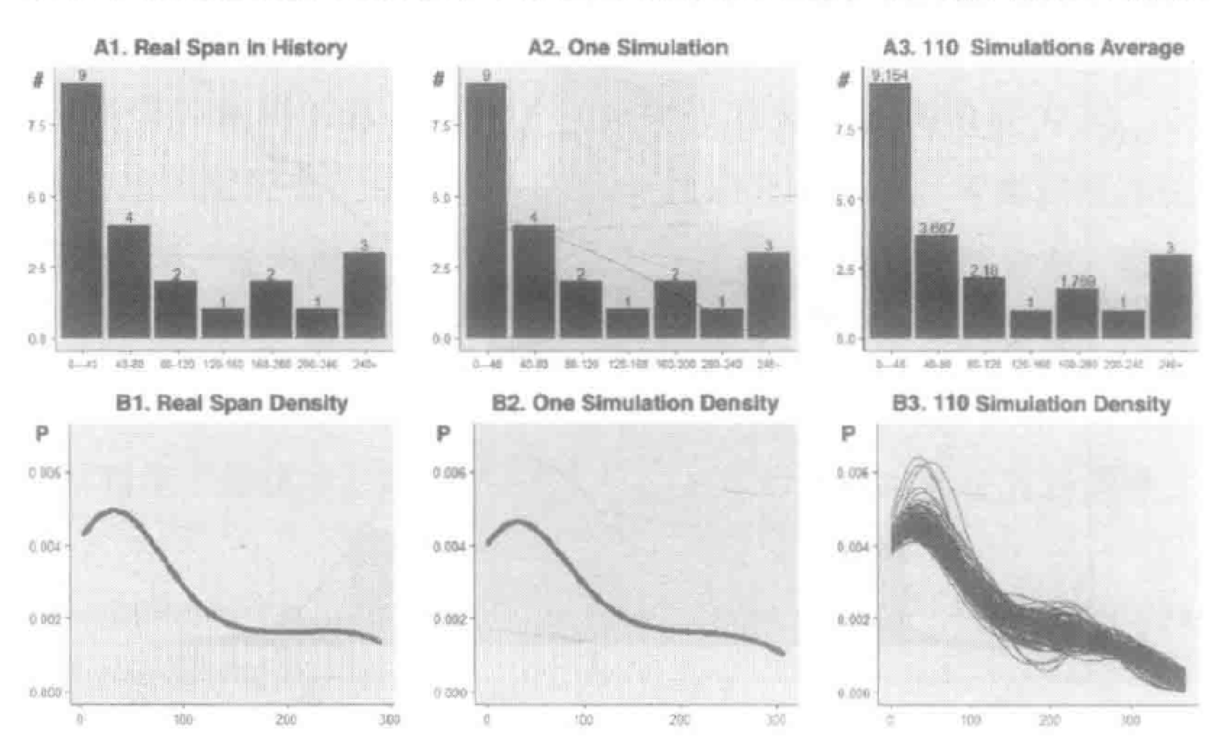


图2　沙堆模型模拟结果的分布拟合

注：子图2A1、2A2和2A3显示了真实帝国、单次最佳模拟结果和100个模拟结果的分布情况。其中，y轴为帝国数量，x轴为帝国国祚的区间范围；子图B1、B2和B3对真实帝国、单次最佳模拟结果和100个模拟结果的核密度函数进行可视化。y轴表示概率，x轴表示帝国寿命即持续时间。

分布［54］。图2证实了阿尔贝斯曼（Arbesman，2011）结论，帝国寿命分布遵循无记忆指数分布［14］。概率密度图进一步检验分布特征。帝国寿命序列的总体趋势可以绘制成图2中B1、B2和B3的平滑曲线，呈现右长尾分布。图A2与B1的真实寿命的平滑概率密度近似符合指数分布。图A2与B2中的最佳模拟结果和图A3、B3中的多次模拟结果（N=110）的平滑曲线也近似符合指数分布。

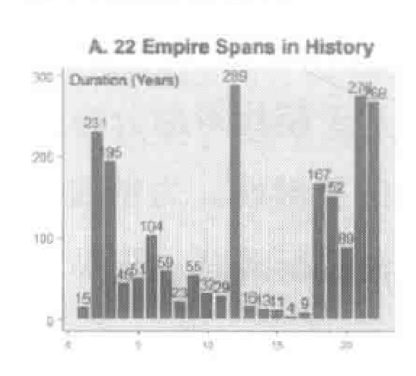
（二）整好22个朝代寿命的精确匹配

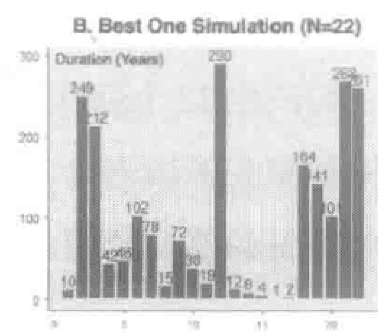
我们逐一比较模拟获得的历史中的朝代寿命。首先从1000次模拟中选择110次结果（均产生22个朝代）。基于朝代寿命分布拟合，可以进一步验证沙堆模型在帝国寿命方面的稳健性。为精准拟合朝代寿命，我们需要对历史和模拟帝国的寿命进行排序。获得22对（模拟获得+真实历史）朝代，两个寿命彼此最为接近。最后，我们恢复真实的历史顺序，模拟朝代的顺序随之更新。为满足真实历史中2133年的总历史时间，我们需要对帝国寿命进行优化，消除重叠的年份。如元朝建立于1271年，而宋朝灭亡于1279年，因此用1279年计算元朝寿命。但是从整体而言，这种微观操作对整个大历史不存在显著性影响。

图3A展示了从秦朝到清朝的22个王朝的历史进程。顺序如下：秦朝（15年）、西汉（231年）、东汉（195年）、三国时期（45年）、西晋（51年）、东晋（104年）、刘宋（59年）、南齐（23年）、南梁（55年）、南陈（32年）、隋朝（29年）、唐朝（289年）、后梁（16年）、后唐（13年）、后晋（11年）、后汉（4年）、后周（9年）、北宋（167年）、南宋（152年）、元朝（89年）、明朝（276年）、清朝（268年）。22个朝代中，唐朝寿命最长，长达289年。五代十国

中的后汉仅有4年，是历史上最短的一个朝代。

图3B展示了根据最优一次模拟的结果所呈现的大历史过程（秦汉到明清），显然与中国2133年的历史吻合较好。基于一对一匹配型考察，模拟结果能够较好地还原历史。（1）历史趋势拟合。汉唐之间，有很多朝代跨度很短。在历史后半段，有多个大朝代寿命超过100年。（2）一对一帝国寿命拟合。模拟与真实寿命之间误差很小，一般在10年以内。图3C展示100次模拟结果，能够较好地还原历史。





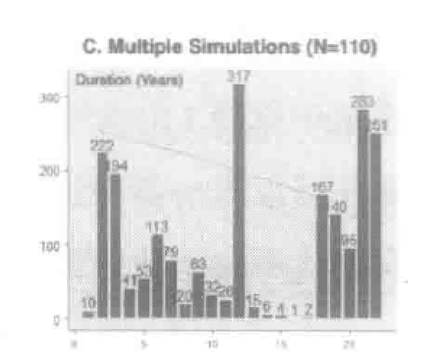


图3 考察22个朝代寿命的历史拟合度

注：我们比较真实历史和模拟朝代的寿命。在每个子图，x轴表示历史朝代顺序，y轴表示寿命（年）。图3A对朝代寿命按照历史顺序进行可视化；图3B展示最佳一次模拟的寿命；图3C展示100次模拟的平均寿命。

（三）少于22个朝代的寿命匹配

在最优解1000次模拟结果中（图1），还存在很大比例的非22朝代情况。为进一步验证沙堆模型稳健性，增强ABM仿真结果的可解释性，我们放宽了朝代数量范围。探索在其他数量的历史情境下，朝代寿命之间的匹配度。从历史本身的角度出发，这种探索也是有意义的。不同的历史学家，其视角或标准存在差异，朝代划分本身就存在一定程度的争议性，尤其是对一些动荡时期、短命朝廷（低于10年的小朝廷）的处理，一直以来就存在随机性、主观性、

模糊性。因此，我们一方面考察低于22个的情况（如19、20、21)、同时也考察高于22个的情况（如23、24、25)。计算朝代寿命的匹配度方面，依然采用公式（2)。本部分，考察19、20、21个朝代寿命的仿真匹配度。具体结果如下。

第一次最优解1000次模拟中，有70次涌现19个朝代。

在很大程度上，可用19个朝代（附表2）刻画中国大历史。我们暂时忽略3个处在动荡期的短命王朝，即后汉（4年)、后周（9年)、后晋（11年)。我们比较了19个模拟朝代和历史朝代的寿命，以验证拟合度。图4中的A1呈现包括19个主要朝代帝国的中国大历史。A2显示70次模拟中最好的一次模拟结果。总的来说，仿真模拟的吻合度极高。对于寿命较长的帝国，模拟吻合度更高。例如第二（西汉)、第三（东汉）朝代持续时间长达231、195年，与模拟寿命237、209年非常接近（差距5年、14年)；寿命最长的是持续289年的第十二朝代，模拟寿命305年，差距（16年）较小；后三个朝代寿命是89年（元)、276年（明）和268年（清)，模拟寿命87、281、262年，差距极小（2、5、6年)。A3针对19组朝代寿命拟合误差进行可视化，发现整体趋势为零。我们还考察了多次模拟均值结果（N=70）的拟合性。如A4，多次模拟均值也符合真实历史。例如：第二、第三个朝代寿命为221年、197年，仿真寿命231年、195年，误差仅为10、2年。最后三个朝代有99年、290年、257年，真实历史寿命为89年、276年、268年，误差仅为10、19、11年。A5中可视化的误差趋势，平滑曲线在零均值附近上下波动。

第二次最优解1000次模拟中，有100次涌现20个朝代。

我们忽略2个“短命朝廷”，即后汉（4年)、后周（9年)，保留后晋（11年)。表2中20个朝代，也能够在很大程度上代表2132

年的大历史。我们构建20组仿真-历史朝代的寿命值，并计算和考察拟合度。图4中的B1将秦汉到明清的20个朝代按照历史顺序进行可视化。B2显示了最佳的一次模拟结果，总的来看，很符合现实。对于寿命较长的帝国，模拟结果非常吻合。例如，第二个朝代（西汉）仿真寿命为235年，历史是231年，误差仅为4年；最后两个模拟朝代寿命分别为273、263年，历史寿命分别为276、268年，误差仅为3、5年。B3进一步考察误差趋势，平均值为零。除了单次最佳模拟，我们考察100次模拟结果的拟合度。B4表明，历史拟合度极高。例如，第二、第三个模拟朝代持续226、194年，真实历史是231、195年，误差仅为5、1年；最后三个模拟朝代寿命是96、289、255年，历史中寿命为89、276、268年，误差仅为7、13、13年。历史与模拟寿命之间的吻合度极高。B5对100次模拟结果的拟合度进行可视化，总体接近零水平线。

第三次最优解1000次模拟中，有102次涌现21个朝代。

仅忽略一个朝代，即后汉（4年）。表2列出的21个帝国，在更大程度上代表大历史。图4中的C1按照历史顺序对21个朝代寿命进行可视化。C2展示了最佳的一次模拟结果。吻合度极高，特别是对长寿命者。比如第二个、第三个仿真朝代的寿命为245、203年，误差仅有14、8年。最后三个仿真寿命为88、269、253年，误差仅为1、7、15年。C3表明，误差在零值附近波动。C4呈现102次模拟结果，更拟合历史。例如，第二、第三个仿真朝代寿命为230、198年，历史国祚为231、195年，误差仅为1、3年；最后五个仿真朝代的寿命为167（北宋）、152（南宋）、89（元）、276（明）、268（清）年。模拟寿命为170、142、95、285、257年，误差仅为3、10、6、9、9年。C5表明，多次模拟均值的寿命误差趋势为零均值。

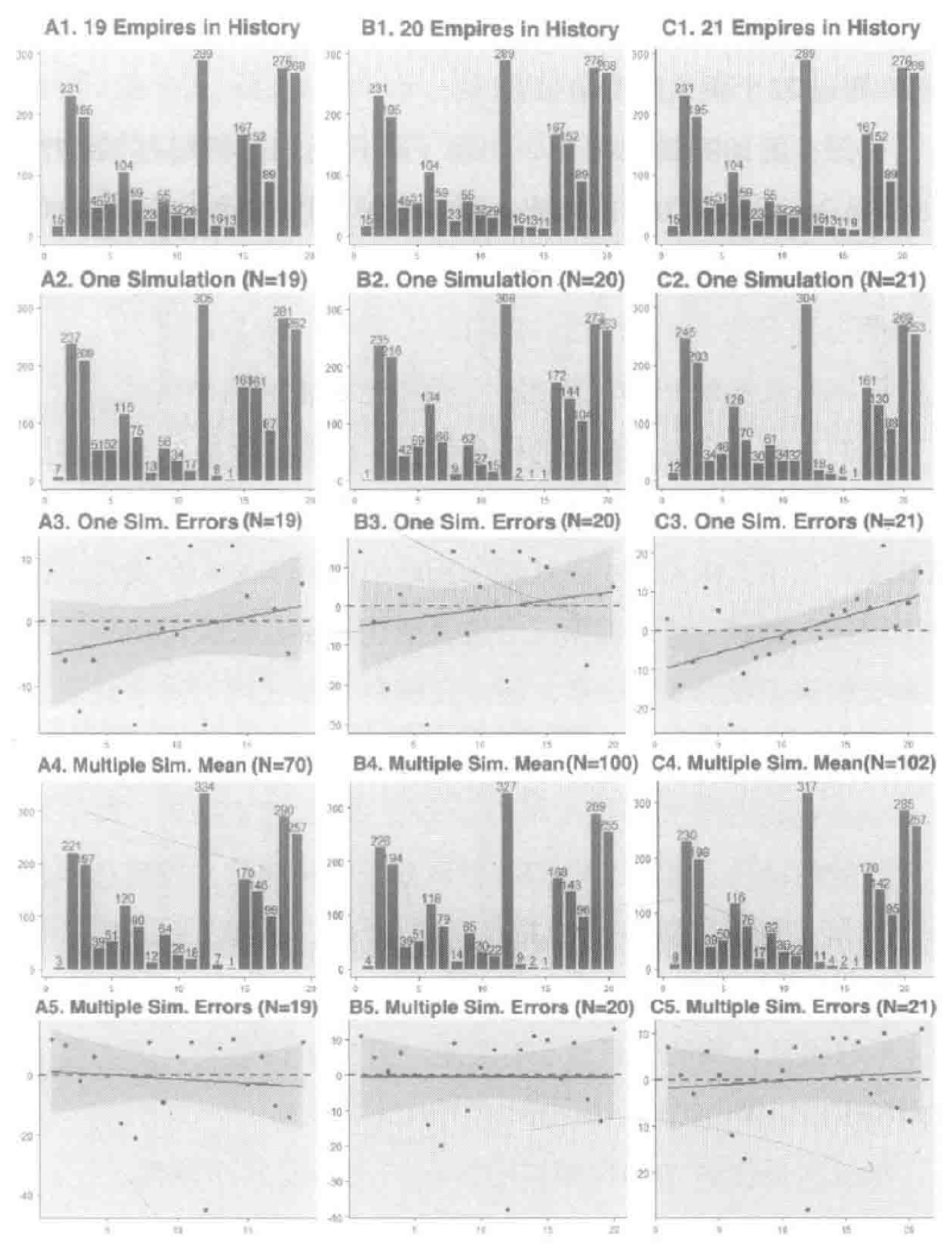


图4 低于22个帝国的匹配

注：图4中的A1—A5验证仿真拟合度；B1—B5对20个误差值进行可视化；C1—C5验证模拟的拟合度。对于朝代数量19、20和21，使用最佳单次模拟、多次模拟均值进行拟合度验证，并绘制A3、B3和C3以及A5、B5和C5中的误差图。每个子图中，x轴表示朝代的历史顺序，y轴表示持续时间或误差。

(四) 超过22个朝代的寿命拟合

如图1在1000次模拟中，出现了朝代涌现数量超过22的情况(如23、24、25)。为匹配历史，需要去掉1个、2个、3个寿命最短的仿真朝代，获得22对历史-仿真寿命值。因为23~25仿真朝代中，寿命最小者不到5年（1、2、4年）。

第四次最优解1000次模拟中，有107次涌现23个朝代。

我们去掉一个寿命最小的仿真朝代（1年），保持22个真实历史朝代数量（表示为23→22过程）。图3中的A1展示了最佳最佳模拟的效果，总体上误差最小。比较图3的A1和图5的A1，发现总体拟合度较高。例如，第二、第三个仿真朝代寿命为226、213年，与真实历史的误差仅为5、18年，误差比例在10%之内。对于北宋，模拟寿命是165年，历史是167年，误差仅为2年。如图5的A2所示，22个误差值的趋势在零值范围内。基于107次模拟结果，取平均值作为稳健性结果。如图5的A3，对误差进行可视化，发现总体趋势为零均值。例如，第一个模拟朝代是12年，误差3年；第二个是221年，误差10年；第三个是190年，误差仅为5年；模拟元朝时仅有4年的误差；明朝也只有6年误差。图5的A4表明，多次模拟结果是稳健的（N=107）。

第五次最优解1000次模拟中，有89次涌现24个朝代。

我们忽略寿命最短的2个模拟朝代（1年与2年），2132年大历史仍然被视为一个由22个帝国组成的连续体（表示为24→22）。通过图3的A中的真实历史数据比较，考察模型研究的稳健性（鲁棒性）。基于图5的B的最佳模拟结果发现，拟合度得到坚实支持。例如，第一个模拟朝代仅有5年误差，第二个和第三个朝代有4年误

差。对于最长的朝代（唐），误差仅为7年（=289–282）。对第十八个朝代（北宋），误差仅为1年。第二十（元朝）、第二十一（明朝）个朝代模拟误差仅为3年、5年。图5中的B2中绘制了误差趋势，在

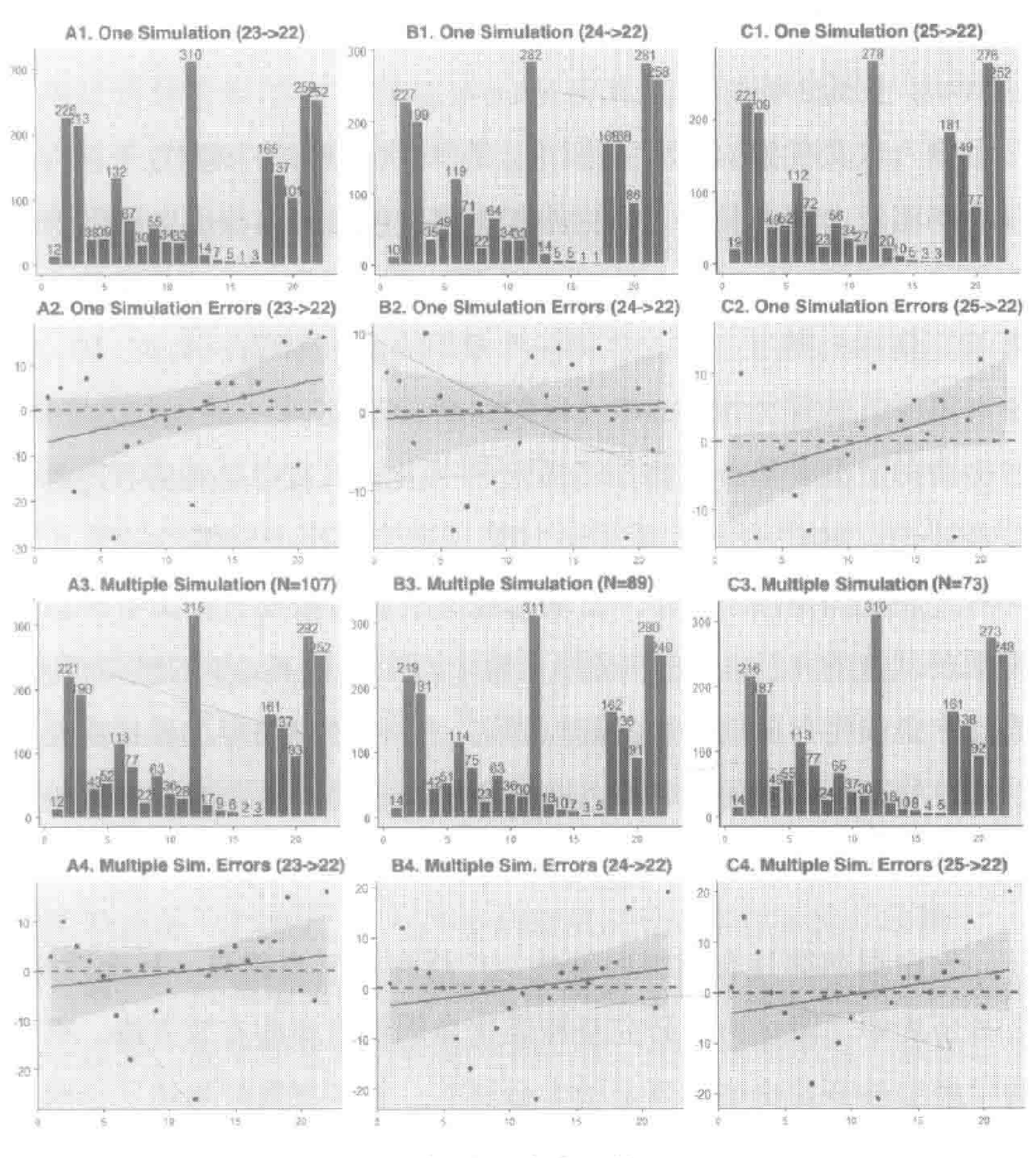


图5　超过22个帝国的匹配

注：图5中的A1—A4考察23个朝代的拟合度；B1—B4考察24个朝代的拟合度；图C1—C4考察25个朝代的拟合度。对23、24和25个朝代，分别考察单次最佳模拟、多次模拟结果的匹配度。我们绘制子图A2、B2和C2以及A4、B4和C4的误差。在每个子图中，x轴表示朝代的历史顺序，y轴表示持续时间或误差。

零均值附近波动。我们考察89次模拟的稳健结果，如B4所示，误差似乎也接近于零。B3按照历史顺序，展示了模拟寿命，拟合度良好。第一个模拟朝代的误差为1年，第三个误差为8年。第十八（北宋）误差为5年。第二十（元）、第二十一（明）个朝代的误差同样很微小，仅为2年和4年。

第六次最优解1000次模拟中，有73次涌现25个朝代。

忽略三个寿命低于5年的仿真朝代，可以获得与历史匹配的22个朝代（图5中的C1）。比较图1的A1、图5的C1，发现历史吻合度极高。例如，第一、第二、第三个朝代的模拟误差仅为4、10、14年。第十九的误差仅为3年（=152−149），第二十一的误差为零（=276−276）。图5中的C2将21个误差可视化，发现总体趋势在零均值附近波动。如C3所示，我们还计算了多次模拟的稳健性结果（N=73）。与图1中的A1比较，发现拟合度极高。第一、第二、第三个模拟朝代仅有4、15、8年误差。第十八个朝代有6年的模拟误差，第二十个朝代有6年的模拟误差，第二十一个仅有3年的模拟误差。

六、结论和讨论

生命周期律具有超级稳健性，广泛存在于自然界、社会系统、人类历史中。量化历史研究是探索生命周期规律的有效方法，很多数学模型被开发出来解释帝国生命周期。如塔塞佩拉阐述了帝国兴起和衰落的数学模型（Taagepera，1978）；赫马拉泽提出了罗马皇帝的指数分布（Khmaladze et al.，2007）；阿尔贝斯曼证明了帝国的无记忆指数分布（Arbesman，2011）；马尔凯蒂和奥苏贝尔以领土扩张为基本参数，以税收、公共支出、腐败为主要指标，开发了数

学模型（Marchetti & Ausubel，2012）。在中国，也有学者引入了随机增长模型分析王朝周期律（Chan & Laffargue，2016）。现有的量化历史研究，提供了必要的参数借鉴与模型参考。但是，也存在动态性不足的问题。本文的创新性在于超越了量化历史研究的传统，使用计算社会科学（多智能体ABM仿真）方法研究历史动力学。我们基于自组织临界性（沙堆模型）研究帝国兴衰动力学，揭示了自然系统、社会系统之间的跨系统相似性。研究表明，帝国社会是一个典型的、原理平衡态的耗散系统。

第一、社会系统、物理系统具有同构性。自然系统和社会系统有相似之处（Barnes & Wilson，2014），本文的沙堆模型与帝国体系的相似性获得了验证。首先，沙子累积过程与帝国分层结构内的社会压力积累有相似之处。累积过程的可以是线性或非线性的。其次，沙堆和帝国会自动到达临界状态。作为一个包含许多组件及其相互作用和关系的复杂系统，我们找到了简单有效的仿真模型来刻画人类帝国社会系统的本质特征。沙堆达到临界状态前，沙粒处于短期平衡状态。随着加入的沙粒越多，到达临界阈值时，可能导致大规模的崩溃甚至一系列的连锁反应，类似于帝国的混乱陨落。当社会问题积累到一定程度，一场小危机就可能导致现存帝国的崩溃。

第二、人类历史中，帝国生命周期一直存在。正如个人的出生、成长、成年和衰老的生命周期一样，帝国也有出生、成长、兴盛和衰亡的生命周期。此外，公司、行业、谣言和感染也有相似的生命周期模式（Rosecrance，1987；Modelski & Thompson，1988）。对于中国和世界上所有的帝国来说，帝国的兴衰肯定会发生，不受任何其他因素的影响。历史上，旧帝国的灭亡总是伴随着新帝国的崛起。因此，旧帝国被新帝国取代的历史趋势也是不可避免的。无

论统治者采取什么样的措施，有什么样的初始条件，历史上的帝国最终都会自动陨落、消失。历史上帝国坚实的生命周期模式是持久和稳定的，因此可以用自组织临界性来解释，并通过沙堆模型来模拟和研究。

第三，帝国兴衰动力学遵循自组织临界性规则。每个帝国都有特定的寿命。随着帝国的仿真壮大，社会结构性问题的严重性也开始不断加深，社会矛盾逐渐积累，最终削弱国本，导致确定性崩溃（Butzer，2012）。当整个社会处于完全混乱状态时，国家是难以控制风险的。由于社会问题仍未解决，持续的朝代更迭现象将持续存在。所以帝国有生命周期，旧帝国被新帝国取代是必然的。对于自组织临界性，沙堆模型在固定参数下，捕捉了帝国自动演化到临界状态的过程。系统的演化是各要素的相互作用，从而产生了自组织临界性。帝国体系还受到社会分类、财政收入、外来威胁、人口增长、耕地、气候等因素及其相互作用的影响（Collins，1995；Weiss & Bradley，2001）。

第四，沙堆模型实现了历史过程反演，并具有预测作用。我们使用多种计算拟合度的指标（标准）寻找最优解，并验证沙堆模型的稳健性，实现了高精度、多场景的精准匹配。第一，历史朝代数量实现100%匹配。基于最优解，重复1000次模拟，分布接近正态分布，平均值为22，完全吻合真实历史。其次，朝代寿命分布实现高精度匹配。历史上，帝国寿命呈现右长尾分布。许多朝代的寿命短于40年，同时也存在长寿的朝代。仿真也获得了同样的朝代寿命的右长尾分布。分布的概率密度函数也实现了精准匹配。历史与模拟的朝代寿命，都遵循幂律分布。我们验证了模型的扩展性、可解释性，在低于（19、20、21）和高于（23、24、25）22朝代模拟

的情况下，沙堆模型仍然具有令人满意的拟合度和解释力。

第五，多次模拟的均值结果优于单次最佳模拟。在最优解1000次模拟中，我们对19、20、21、22、23、24和25个朝代的情况，都考察了多次模拟结果（取平均值），以便于比较单次最佳模拟结果、多次模拟均值结果。研究发现，多次模拟的稳健性比单次最佳模拟显然更优。图2中，我们将朝代寿命多次模拟的均值结果进行可视化；比较图4中的A3—A5、B3—B5和C3—C5，发现多次模拟均值的误差分布特征，比单次最佳模拟结果更接近零均值趋势；在图5中也非常明显，比较图5中的A2—A4和C2—C4，多次模拟比单次最佳模拟具有更多的零误差值。因此，多次模拟均值比单次最佳模拟结果具有更稳健的拟合度效果。

第六，多智能体ABM仿真具有研究优势。中国古人使用五行学说（五德终始说）等解释王朝更替（Chen，2014），但非科学的思维方式。也有人认为王朝周期律主要是由气候变化引起（Fan，2010），这也是本末倒置的片面观点。莫德尔斯基发现帝国海洋霸权符合“长周期理论”（Modelski & Thompson，1988），即每一个帝国或霸权统治了一定时期之后，都注定要崩溃。生命周期分为全球性战争、全球霸权出现、去合法化、权力分散等四个阶段，在国际关系方面具有解释度（Rosecrance，1987），但仍然是高级描述，没有揭示内在本质。巴菲尔德认为，中原王朝和游牧民族之间具有周期性关系，如汉族和匈奴、唐朝和突厥、宋和蒙古人、明和后金的循环（Barfield，1991），其能够解释帝国兴衰，但“隔靴搔痒”，缺乏内在机制的探讨。我们基于自组织临界性，将自然系统、社会系统纳入统一的考察框架，无疑提升了社会历史研究的科学性、精确性。

七、局限和未来展望

尽管自组织临界性（沙堆模型）可以解释帝国动力学，但更加高颗粒度的历史过程仍然超出了建模与模拟的能力。所采取的方法，首先是确定历史顺序，然后相应地匹配真实历史寿命，尚未实现22个朝代“一笔画”。此外，历史上有不同版本的朝代划分（22、23、24个），这给模型带来复杂性挑战。本模型可用于其他社会系统研究，如组织、企业、疾病、谣言等，这些领域也都符合自组织临界性与生命周期特征。未来研究中，我们将从中国王朝周期律扩展到全球历史中的帝国。我们还将拼接出一个基于沙堆模型的“世界体系”。每个国家都是一个沙堆，整个世界是一个沙堆系统。局部的沙堆演化最终导致整个世界的崩溃。

参考文献

Arbesman, Samuel 2011, "The life-spans of empires." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 44, 127-129.

Babcock, K. & Westervelt R. M. 1990, "Avalanches and self-organization in cellular magnetic-domain patterns." *Physical review letters* 64(18), 2168.

Bak, P. 1991, "Self-organized critically." *Scientific American* 264, 46-53.

—2013, "How nature works: the science of self-organized criticality." *Springer Science & Business Media*.

— & Chao Tang 1989, "Earthquakes as a self-organized critical phenomenon." *Journal of Geophysical Research: Solid Earth* 94(B11), 15635-15637.

— & Chao Tang, Kurt Wiesenfeld 1987, "Self-organized criticality: An explanation of the 1/f noise." *Physical review letters* 59(4), 381.

— & Chao Tang, Kurt Wiesenfeld 1988, "Self-organized criticality." *Physical review A* 38(1), 364.

Barfield, J. Thomas 1991, "Inner Asia and cycles of power in China's imperial dynastic history." *Rulers from the Steppe: State Formation on the Eurasian Periphery* 21–62.

Barkawi, Tarak & Mark Laffey 2002, "Retrieving the imperial: empire and international relations." Millennium – *Journal of International Studies* 31(1), 109–127.

Barnes, Trevor J. & Matthew W. Wilson 2014, "Big data, social physics, and spatial analysis: the early years." *Hydrotechnical Construction* 1(1), 613–616.

Basu, Tapan 2012, "Caste Matters: Rabindranath Tagore's Engagement with India's Ancient Social Hierarchies." *South Asia: Journal of South Asian Studies* 35(1), 162–171.

Butzer, Karl W. 2012, "Collapse, environment, and society." *Proceedings of the National Academy of Sciences of the United States of America* 109(10), 3632–3639.

Chan, Kenneth S. & Jean-Pierre Laffargue 2016, "A dynamic model of taxation, corruption, and public investment in the dynastic cycle: The case of Imperial China." *Macroeconomic Dynamics* 20(8), 2123–2147.

Chase-Dunn, C. & Alexis Alvarez, Hiroko Inoue, Richard Niemeyer, Anders Carlson, Ben Fierro, Kirk Lawrence 2006, "Upward Sweeps of Empire and City Growth Since the Bronze Age." *The Institute for Research on World-Systems.*

Chase-Dunn, C., Richard Niemeyer, Alexis Alvarez, Hiroko Inoue, Kirk Lawrence & James Love 2010, "Cycles of Rise and Fall, Upsweeps and Collapses: Changes in the scale of settlements and polities since the Bronze Age." *History and Mathematics: Processes and Models of Global Dynamics* 64–91.

Chase-Dunn, C. 1990, "World-state formation: historical processes and emergent necessity." *Political Geography Quarterly* 9(2), 108–130.

Chen, Yuan Julian 2014, "Legitimation discourse and the theory of the five elements in imperial China." *Journal of Song-Yuan Studies* 44, 325–364.

Cheon, Taksu & Sergey S. Poghosyan 2017, "Spiral orbits and oscillations in historical evolution of empires." *Physica A: Statistical Mechanics and its Applications* 469, 353–362.

Collins, Randall 1995, "Prediction in macrosociology: The case of the Soviet col-

lapse." *American journal of sociology* 100(6), 1552–1593.

de Quiroga, Pedro López Barja 2010, "Empire Sociology: Italian freedmen, from success to oblivion." *Historia: Zeitschrift für Alte Geschichte* 321–341.

Diodati, P., F. Marchesoni & S. Piazza 1991, "Acoustic emission from volcanic rocks: an example of self-organized criticality." *Physical Review Letters* 67, 2239.

Eckhardt, William 1990, "Civilizations, empires, and wars." *Journal of Peace Research* 27, 9–24.

Elmer, Franz-Josef 1997, "Self-organized criticality with complex scaling exponents in the train model." *Physical Review* E 56, R6225.

Fan, Ka-wai 2010, "Climatic change and dynastic cycles in Chinese history: a review essay." *Climatic Change* 101, 565–573.

Ferguson, Niall 2006, "Empires with expiration dates." Foreign Policy 46–52.

—2010, "Complexity and collapse: empires on the edge of chaos." *Foreign affairs* 89 (2), 18–32.

Garrido, Pablo, Shaun Lovejoy & Daniel Schertzer 1996, "Multifractal processes and self-organized criticality in the large-scale structure of the universe." *Physica A: Statistical Mechanics and its Applications* 225, 294–311.

Gerald, Segal & Kennedy Paul 1988, "The rise and fall of the great powers: economic change and military conflict from 1500 to 2000." *International Affairs (Royal Institute of International Affairs* 1944–)

Gudner, Andre & Barry K. Gills 1992, "World system cycles, crises, and hegemonial shifts, 1700 BC to 1700 AD." *Review (Fernand Braudel Center)* 621–687.

Hall, Thomas D. & Christopher Chase-Dunn 1993. "The world-systems perspective and archaeology: Forward into the past." *Journal of Archaeological Research* 1, 121–143.

Held, G. A., D. Solina, H. Solina, D. Keane, W. Haag, P. Horn & G. Grinstein 1990, "Experimental study of critical-mass fluctuations in an evolving sandpile." *Physical Review Letters* 65, 1120.

Iberall, Arthur S. 1987, "A physics for studies of civilization." *Self-Organizing Systems*. Springer 521–540.

—1984, "Contributions to a physical science for the study of civilization." *Journal of Social and Biological Structures* 7, 259–283.

Jiang, Zhaohui 2007, "Cultural Root of Power Hierarchy." *Journal of Guangxi Normal University (Philosophy and Social Sciences Edition)*.

Khmaladze, Estate, Ray Brownrigg & John Haywood 2007, "Brittle power: On Roman Emperors and exponential lengths of rule." *Statistics & probability letters* 77, 1248–1257.

Liu, Li 1996, "Mortuary ritual and social hierarchy in the Longshan culture." *Early China* 21, 1–46.

Mahaffy, John Pentland 1899, "A history of Egypt under the Ptolemaic dynasty." *Methuen & Company*.

Malamud, Bruce D., Gleb Morein & Donald L. Turcotte 1998, "Forest fires: an example of self-organized critical behavior." *Science* 281, 1840–1842.

Marchetti, Cesare & Jesse Ausubel 2012, "Quantitative dynamics of human empires." *International Journal of Anthropology* 27, 1–62.

Modelski, George & William R. Thompson 1988, "The Long Cycle of World Leadership." *Seapower in Global Politics*, 1494 - 1993 97–132.

Norkus, Zenonas 2008, "The Grand Duchy of Lithuania in the retrospective of comparative historical sociology of empires." *World Political Science* 3.

Picoli, Sergio, Marcelo del Castillo–Mussot, Haroldo V. Ribeiro, Ervin K. Lenzi & Renio S. Mendes 2014, "Universal bursty behaviour in human violent conflicts." *Scientific reports* 4, 1–3.

Portugali, Juval 1997, "Self–organizing cities." *Futures* 29, 353–380.

Roberts, D. C. & D. L. Turcotte 1998, "Fractality and self–organized criticality of wars." *Fractals* 6, 351–357.

Rosecrance, Richard 1987, "Long cycle theory and international relations." *International organization* 41, 283–301.

Sapozhnikov, Victor B. & Efi Foufoula-Georgiou 1997, "Experimental evidence of dynamic scaling and indications of self-organized criticality in braided rivers." *Water Resources Research* 33, 1983–1991.

—1995, "The archaeology of empires: a view from South Asia." *Bulletin of the American Schools of Oriental Research* 299, 3–11.

Shimada, Ichiro & Tomio Koyama 2015, "A Theory for Complex Systems Social Change: An Application of a General 'Criticality' Model." *Interdisciplinary Descrip-*

tion of Complex Systems 13(3), 342–353.

Sinopoli, Carla M. 1994, "The archaeology of empires". *Annual Review of Anthropology* 23, 159–180.

Stauffer, Dietrich. & Didier Sornette 1999, "Self-organized percolation model for stock market fluctuations." *Physica A: Statistical Mechanics and its Applications* 271, 496–506.

Steinmetz, George 2016, "Social fields, subfields and social spaces at the scale of empires: Explaining the colonial state and colonial sociology." *The Sociological Review* 64, 98–123.

Taagepera, Rein 1978a, "Size and duration of empires growth-decline curves, 3000 to 600 bc." *Social Science Research* 7, 180–196.

—1978b, "Size and duration of empires: systematics of size." *Social Science Research* 7(2), 108–127.

—1979, "Size and duration of empires: growth-decline curves, 600 BC to 600 AD." *Social Science History* 3, 115–138.

Triandafyllidou, Anna & Mariangela Veikou 2002, "The hierarchy of Greekness: Ethnic and national identity considerations in Greek immigration policy." *Ethnicities* 2, 189–208.

Turchin, Peter 2005, "Dynamical feedbacks between population growth and sociopolitical instability in agrarian states." *Structure and Dynamics* 1.

Weiss, Harvey & Raymond. S Bradley 2001, "What drives societal collapse?" *Science* 291, 609–610.

Wimmer, Andreas & Min Brian 2006, "From empire to nation-state: Explaining wars in the modern world, 1816 - 2001." *American sociological review* 71, 867–897.

Yao, Alice 2005, "Scratching beneath iconographic and textual clues: A reconsideration of the social hierarchy in the Dian culture of Southwestern China." *Journal of Anthropological Archaeology* 24, 378–405.

Yegorov, Yuri, Dieter Grass, Magda Mirescu, Gustav Feichtinger & Franz Wirl 2020, "Growth and collapse of empires: a dynamic optimization model." *Journal of Optimization Theory and Applications* 186, 620–643.

Yuan, Jian, Yong Ren & Xiuming Shan 2000, "Self-organized criticality in a computer network model." *Physical Review* E 61, 1067.

Zhukov, Dmitry S., Valery V. Kanishchev & Sergey K. Lyamin 2016, "Application of the theory of self-organized criticality to the investigation of historical processes." *Sage Open* 6(4).

附录:

附表1:中国历史上的22个帝国

帝国或王朝	诞生年份	结束年份	寿命(Y_i,年)
秦	公元前221	公元前206	15
西汉	公元前206	公元25	231
东汉	公元25	公元220	195
三国	公元220	公元265	45
西晋	公元265	公元316	51
东晋	公元316	公元420	104
南北朝—宋	公元420	公元479	59
南北朝—齐	公元479	公元502	23
南北朝梁	公元502	公元557	55
南北朝—陈	公元557	公元589	32
隋	公元581	公元618	37
唐	公元618	公元907	289
五代十国—后梁	公元907	公元923	16
五代十国—后唐	公元923	公元936	13
五代十国—后晋	公元936	公元947	11
五代十国—后汉	公元947	公元951	4
五代十国—后周	公元951	公元960	9
北宋	公元960	公元1127	167
南宋	公元1127	公元1279	152
元	公元1271	公元1368	89
明	公元1368	公元1644	276
清	公元1636	公元1912	268

注:由于史料不同,各个朝代起止时间众说纷纭。本研究中,各朝代起止时间的主要参考文献为《中国通史》(吕思勉,2012)和《中国历史纪年表》(方诗铭,

2013）。但是，为保障朝代起止时间连续性，本研究对史学界中有争议的部分朝代起止时间进行了调整。因此，本研究与其他研究文献略有出入，该偏差不影响研究结论。

附表2：中国历史19—21个帝国模拟结果

19帝国	国祚（历史）	最佳模拟	最佳模拟误差	多次模拟均值（N=70）	均值误差
秦	15	7	8	3	12
西汉	231	237	−6	221	10
东汉	195	209	−14	197	−2
三国	45	51	−6	39	6
西晋	51	52	−1	51	0
东晋	104	115	−11	120	−16
宋	59	75	−16	80	−21
齐	23	13	10	12	11
梁	55	56	−1	64	−9
陈	32	34	−2	26	6
19帝国	国祚（历史）	最佳模拟	最佳模拟误差	多次模拟均值（N=70）	均值误差
隋	29	17	12	18	11
唐	289	305	−16	334	−45
后梁	16	8	8	7	9
后唐	13	1	12	1	12
北宋	167	163	4	170	−3
南宋	152	161	−9	146	6
元	89	87	2	99	−10
明	276	281	−5	290	−14
清	268	262	6	257	11
20帝国	国祚（历史）	最佳模拟	最佳模拟误差	多次模拟均值（N=70）	均值误差
秦	15	1	4	14	11

续表

西汉	231	235	226	-4	5
东汉	195	216	194	-21	1
三国	45	42	39	3	6
西晋	51	59	51	-8	0
东晋	104	134	118	-30	-14
宋	59	66	79	-7	-20
齐	23	9	14	14	9
梁	55	62	65	-7	-10
陈	32	27	30	5	2
隋	29	15	22	14	7
唐	289	308	327	-19	-38
后梁	16	2	9	14	7
后唐	13	1	2	12	11
后晋	11	1	1	10	10
北宋	167	172	168	-5	-1
20帝国	国祚(历史)	最佳模拟	最佳模拟误差	多次模拟均值(N=70)	均值误差
南宋	152	144	143	8	9
元	89	104	96	-15	-7
明	276	273	289	3	-13
清	268	263	5	255	-13
21帝国	国祚(历史)	最佳模拟	最佳模拟误差	多次模拟均值(N=102)	均值误差
秦	15	12	8	3	7
西汉	231	245	230	-14	1
东汉	195	203	198	-8	-3
三国	45	34	39	11	6
西晋	51	46	50	5	1
东晋	104	128	116	-24	-12

续表

宋	59	70	76	−11	−17
齐	23	30	17	−7	6
梁	55	61	62	−6	−7
陈	32	34	30	−2	2
隋	29	32	22	−3	7
唐	289	304	317	−15	−28
后梁	16	18	11	−2	5
后唐	13	9	4	4	9
后晋	11	6	2	5	9
后周	9	1	1	8	8
北宋	167	161	170	6	−3
南宋	152	130	142	22	10
元	89	88	95	1	−6
明	276	269	285	7	−9
清	268	253	257	15	11

附表3:中国历史22—25个朝代模拟结果

王朝持续时间		22帝国		23帝国		23帝国		25帝国	
		最佳模拟	多次模拟均值	最佳模拟	多次模拟均值	最佳模拟	多次模拟均值	最佳模拟	多次模拟均值
秦	15	10	10	12	12	10	14	19	14
西汉	231	249	222	226	221	227	219	221	216
东汉	195	212	194	213	190	199	191	209	187
三国	45	42	41	38	43	35	42	49	45
西晋	51	46	53	39	52	49	51	52	55

续表

王朝持续时间		22帝国		23帝国		23帝国		25帝国	
		最佳模拟	多次模拟均值	最佳模拟	多次模拟均值	最佳模拟	多次模拟均值	最佳模拟	多次模拟均值
东晋	104	102	113	132	113	119	114	112	113
宋	59	78	79	67	77	71	75	72	77
齐	23	15	20	30	22	22	23	23	24
梁	55	72	63	55	63	64	63	56	65
陈	32	38	32	34	36	34	36	34	37
隋	29	19	26	33	28	33	30	27	30
唐	289	290	317	310	315	282	311	278	310
后梁	16	12	15	14	17	14	18	20	18
后唐	13	8	6	7	9	5	10	10	10
后晋	11	4	4	5	6	5	7	5	8
后汉	4	1	1	1	2	1	3	3	4
后周	9	2	2	3	3	1	5	3	5
北宋	167	164	167	165	161	168	162	181	161
南宋	152	141	140	137	137	168	136	149	138
元	89	101	95	101	93	86	91	77	92
明	276	268	283	259	282	281	280	276	273
清	268	261	251	252	252	258	249	252	248

董海军，王圣儒 .E 问卷实务流程探索与趋势展望[M/OL]//赵联飞，赵锋 . 社会研究方法评论：第 2 卷 . 重庆：重庆大学出版社 .

E问卷实务流程探索与趋势展望①

董海军　王圣儒②

摘要：当人类社会进入移动互联网并热议元宇宙时，网络问卷调查业已成为主要的调查研究方法之一。已有研究聚焦于E问卷的特性和运用方向，但在E问卷的具体实务流程上仍涉足不深。本文以一个调查项目为例提炼重点步骤展开案例分析，集中探讨E问卷操作实务中调查计划、调查执行、分析写作等三大流程，初步提出了优化内容设计、完善抽样方法、实施分群驱动等全局改进策略，并展望了E问卷操作实务承办制度化和交互可视化两种发展态势，以期驱动E问卷调查的标准化发展。

关键词：移动互联网+问卷；实务操作；案例分析；标准化

①本研究成果为湖南省智库重大项目（项目编号：20ZWA18）的阶段性成果。

②作者简介：董海军（1978.9–），男，汉族，湖南衡南县人，社会学博士，中南大学公共管理学院副院长，社会学教授、博士生导师，中南大学社会调查与民意研究中心主任、中国社会学会社会调查研究方法专委会主任，主要研究方向：社会学研究方法，社会管理与社会工作。王圣儒，中南大学社会学硕士研究生。联系方式：26079628@qq.com。

Abstract: The online survey has become one of the main survey research methods with the arrival of the mobile internet era and the heated discussion about the Metaverse. The existing studies have focused on the aspects of characteristics and application of E-questionnaire, but there are still some deficiencies in the specific practical process of E-questionnaire. This article extracts the essential steps and performs an analysis of selected cases on the basis of a project, and meanwhile, focuses on the three main processes of E-questionnaire of investigation planning, investigation execution, and analytical writing in the practice of e-questionnaire operation. Three solution improvement strategies, such as optimizing the content design, perfecting the sampling method and implementing the cluster driven are tentatively proposed. Two kinds of tendency, that is contractual institutionalization and visualized interaction, are put forward as development prospects for E-questionnaire operation to promote the standardization.

Key words: Network Questionnaire; Practical Operation; Case Analysis; Standardization

一、引言

移动互联网裂变式的发展（蒋笃军，2021），正深刻厘革着信息传播技术，而调查问卷也同步经历着一场划时代的变革。社会调查的问卷工具已逐渐从传统的纸质问卷，发展成为立足于网络技术的移动互联网+问卷（以下简称E问卷）。E问卷是在移动互联时代下，科技创新与调查实践纵深推进的一种产物（董海军、耿宇，2018），时至今日已迈进全域赋能的新阶段。不论学术界抑或企业，

皆越来越多地依靠其实现提质增效。E问卷的蓬勃发展一方面推动了相关理论稳健阔进，另一方面迎合了时代发展的潮流。

然而，E问卷的实务操作仍存在一些理论争议与实践障碍，这是其公信力尚未得到普遍认可的重要原因之一（邵国松、谢珺，2021）。针对E问卷在实际中的运用，学者们主要从三个角度切入进行研究分析。

第一是E问卷的应用方向。已有研究多是以E问卷的宏观角度作为落脚点。学者们从探测性应用、描述性应用、互动性应用等方面出发（邵国松、谢珺，2021；方佳明、邵培基，2006），探讨了不同行业特征的对象凭借E问卷实现特定统计资料高效率收集及自身吸引力突破等重要问题（张静，2009）。

第二是E问卷的独特优势。国内外研究从实证与理论入手，在回复率、回收分析、反馈时间、调查成本、实践普及等方面，将网络问卷调查同传统纸质问卷调查、电话调查进行比较分析（Tse-Hua & Xitao，2008），得出网络调查兼备时间及费用成本的低廉性、实践的便捷性、内容的智能可塑性、功能的拓展性以及结果获取的即时性等特点（闫建、郑迦元，2020），对E问卷的独特优势及广阔前景构建了详实认知。

第三是E问卷的发展瓶颈。在探讨E问卷价值的相关论述中，学者们梳理了E问卷的发展脉络，反思了其既有的局限和未来发展存在的挑战（邵国松、谢珺，2021）。基于当今社会信息网络化的特性，E问卷主要在调查对象分布不均、硬件系统和软件技术的潜在不可靠性、填写问卷的驱动力不足、调查信度效度低上存在改进空间（陈云松等，2016）。

总体而言，学界对业已普及的E问卷的探究，主要是以实证研

究和局部研究作为载体，聚焦在E问卷的特性和运用，阐释了E问卷为何又如何蓬勃发展，但相对来说，对E问卷的实务操作流程的反思则仍然存在不足。这不仅不利于E问卷的推行，也致使用户承担了不可预估的风险和耗损（李锐、宋铁英，2005）。有鉴于此，本文旨在探索网络问卷调查标准化的实务流程，以期驱动E问卷朝更为规范化和严谨化的方向迈步前行。

二、笃行致远：E问卷实务流程探索的必然性

E问卷的描述性定义为：调查者在社会现实中借助大数据技术，利用终端设备，为实现回收与社会事实休戚相关的调查数据、访谈资料等目标所使用的问卷调查工具（李巍，2009）。在万物互联图景徐徐展开、传统问卷调查研究方式逐渐式微的背景下，问卷调查也围绕数据的流动循环、交互作用，建构成一个日益复杂的E问卷生态系统。越来越融入人们生产生活的网络问卷调查，会被用于持续提升人类福祉，还是被视为一味谋利的商业化工具？能引导E问卷朝着有利于社会发展方向前进的，不仅是技术创新的内在逻辑，更是紧密联系实践的规范流程（王莹、万舒晨，2016）。

第一，探索实务流程是E问卷成熟化的必由之路。遇到E问卷发展中的问题，只有深入实际才能找到答案；面对未来路上的挑战，只有探讨实务才能应付裕如。有了完备的操作实务体系才能提高实务操作能力，提高实务操作能力才能充分发挥操作实务体系的效能（刘慧园，2019）。实现实务与理论的相互促进、循环往复，才能有效推动E问卷成熟发展，做到蹄疾而步稳。

第二，标准化的E问卷实务流程是规范网络调查的重要举措。

互联网的快速发展不仅给社会各领域带来深刻变革，也对问卷调查产生了深远影响，极大地考验着网络问卷调查的规范有序性（刘慧园，2019）。完善网络问卷调查治理要借鉴既有的理论经验，更要以实际为引领，大力推进E问卷操作实务查究（王齐齐，2021）。

第三，规程化的实务流程是E问卷推广开来的行动指南。网络问卷调查工作要推行起来，联系实务是必不可少的。实务操作是直接切入点，也是重要着力点。网络问卷调查工作不能凌虚蹈空，不能脱离实务（刘玉拴，2019）。这是遵从实际、遵循规律的传播。也只有立足实务、联系实际，建立一系列可复制、可借鉴的规程化流程，才能探索出一条符合E问卷自身发展规律的特色推广之路（何哲，2020）。

一言以蔽之，在网络问卷调查论域内探讨操作实务是至关重要的。标准化、规程化的实务流程不仅是E问卷成熟发展的重要宝典，也是规范网络调查的有效帮手，更是贴合社会呼唤的坚实底气。

三、谛分审布：E问卷操作实务流程建构与经验反思

立足于E问卷操作实务的理论缺憾，我们在已有研究基础上，将网络问卷调查操作流程作为考察侧重点，提出改进思路及举措。在实践中，以调查计划阶段、调查执行阶段、分析写作阶段为重点抓手（图1），深耕E问卷实操架构，从具体举措出发织密E问卷实务网。

下文拟结合“Z大学学生兼职情况”调查项目，提炼问卷操作实务三阶段中的核心步骤，对其要领技巧分星擘两；观照现实，紧扣实践需要，出发点是推进网络问卷调查质量工程建设，落脚点是操作实务中可能遇到的问题。

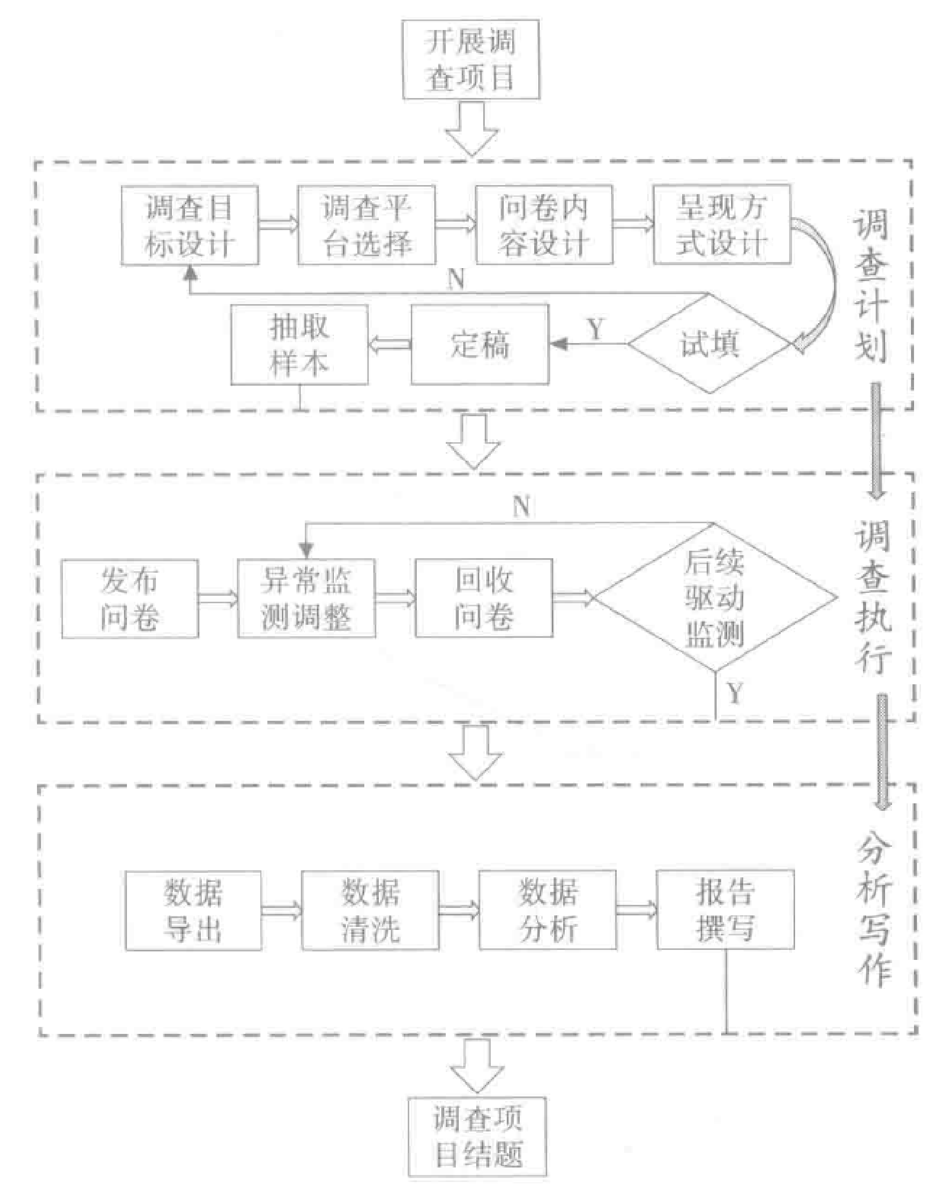


图1　E问卷操作实务流程图

（一）调查计划阶段

这一阶段由目标设计、平台选择、问卷设计、定稿与试填、抽样与确定样本等步骤组成。调查目标是调查项目的信标，任何形式的调查都应先确定调查目标，网络问卷调查也不例外。而后有的放矢地选择调查平台，依赖平台技术开展问卷设计并启动试填写。问卷定稿后，可采取概率抽样和非概率抽样两类方法抽样取本。得当

的抽样方法不仅能从时空场域角度节省人力物力；还可拓展研究结果广度、开掘其深度（Katja et al.，2008）。

1. 问卷设计

问卷设计是E问卷调查计划阶段最重要的步骤，它不是一个静态模式化的阶段，也不是一个千篇一律、故步自封的阶段，而是一个动态的、因地制宜的、与时俱进的过程（阿迪力·努尔，2012）。大数据价值密度低与网络问卷碎片化的结构性特征导致网络调查过程中因填写耐性差而弃填、乱写等行为屡见不鲜；与此同时，由于网络空间匿名性的特质，回收结果与其真实想法存在异质性等现象层见叠出，调查结果信度效度受到影响。对此类困境，可把握调查主题，抓住调查目的根本，从问卷内容和呈现方式两方面锻造完善问卷设计之道进而纾解。

内容设计基本要素主要有三种。（1）封面信设计：封面信是撬动调查对象填写积极性、保障问卷填答质量的有力杠杆（吕正娟，2012）。封面信应署有对调查对象的称呼并阐释调查者身份、调查大致内容和主要目的、调查对象选取方法以及数据使用保密承诺（Bosnjak & Tuten，2003）。（2）问题设计：新一代信息技术与问卷调查的结合，催生了E问卷数字化、信息化、智能化、精简化的理念和模式，在问卷题型设计中发挥重要作用。为提升调查对象满意度，激发实际问卷完成度增长动力，应清理开放性试题比重过大而堆砌的隐性壁垒，将增加封闭式问题比例做量化安排（阿迪力·努尔，2012）。此次调查共设置单项选择题12道、多项选择题8道、排序题1道。在此过程中应兼顾对问卷长度的把控——网络问卷的长度与回复率成负线性相关（吕正娟，2012）。因而在保证问卷设计

“质精”的前提下，当尽可能保持问卷凝练。可采取不显示全部问题数量、题号的设计方式隐藏答题量，为调查对象在填答问卷时提供有效“泄压阀”，继而提高问卷反馈率和应答质量。（3）填答方式设计：网络问卷平台可管制调查对象问卷填答方式。对问卷提前设计填答逻辑勘误、甄别题以将不符合研究要求的样本数据拦截在数据库之外；另外还可依靠限制用户填答工具、问卷填答IP地址，设置选项选择数量、最低答题时长等多重手段复核筛查，控制数据质量（余富强等，2019；Zhifeng et al.，2016）。

移动智能终端应用拓展，勾勒出网络问卷呈现方式设计软件和硬件限制逐步冰解壤分的画卷。网络问卷内容一经确定，便可通过对背景、布局、版式、题号、字体、颜色等设计修缮外观。调查者在设计问卷时，需依照移动智能终端设备特点，注重切合调查对象阅读习惯，遵循统筹技术兼容性和操作便捷性的“友好反馈界面”网络问卷设计原则。在“Z大学学生兼职情况调查”一例中，笔者围绕时间、行业、目的、受骗经历等四个维度，锚定大学生兼职参与基本情况拟定调查。为契合调查对象的阅读惯性和操作方式，问卷选用滚动设计。相较于自动分页式设计，滚动设计中所有问题映现至同一版面，形成高效利用画幅资源的空间格局，切实保障用户对前后语境的理解。此外，单个或者一组问题不宜设置过长，题干、选项、勾选框最好一并呈现，同时可使用加粗字体、设置下划线等方式将题干与选项区别开来。通过排版优化视觉体验，营造有质感的记忆点，避免产生记忆偏差，确保调查对象能抟心揖志填写问卷，提高填写效率、加快进度（卢锦运等，2015）。事实上，目前市面上大多数专业问卷平台皆可做到把脉精准，诊断准确，积极促进所有展示页面、题型显示供给与调查需求耦合，有序融合调查项目特色与个性化设计。

2. 抽取样本

抽取样本包含确定总体和实际抽取两阶段。通常而言，网络问卷调查的目标总体可划分为全体网络用户和部分网络用户。依照调查目标，调查者需对目标总体进行规范，确保其满足同质性、大量性和差异性三个基本特征（卫海英、杨智勤，2002）。根据调查目标确定总体后，可采用概率抽样和非概率抽样两类方法选取合适样本。随着我国互联网渗透率大幅提高（刘洋滔，2021），现有网络调查把非概率抽样作为主要的抽样方式（Mick，1997）。然而，因网络覆盖率和个体差异，偶遇抽样、滚雪球抽样等非概率抽样法带来了网络问卷同温层滚动问题，产生网络问卷调查涵盖误差，严重威胁样本代表性。为解决此类问题，推荐以下三种抽样方式。

（1）在面临一些分布集聚、流动性较小的群体时，可采取与传统调查相似的线下抽样模式。本次“Z大学学生兼职情况调查”项目样本目标明确、边界清晰，即Z大学学生，笔者按照Z大学对专业学科门类的划分，依据10个专业门类以班级为单位对样本进行区分，确保样本具有代表性。使用随机抽样法确定调查班级后，联系所抽取班级施行线上调研。除上述方法外，调查者在线下完成抽样并携带移动设备进行实地调研也为可行之策。此法区别于传统问卷调查，填写过程中移动终端替代纸笔。在监测答题、降低样本涵盖误差的同时，依旧保留电子问卷填答便捷、数据上传即时等优势（张青、瞿莉，2020）。

（2）基于网络地址的随机抽样技术。该种线上抽样方法能有效解决网络问卷同温层滚动问题。鉴于在互联网中，计算机IP地址具有唯一性，因此可以利用随机IP发生软件，在设定IP地址码范围内抽取IP地址，再通过发送电子邮件等方式邀请被调查者填写问

卷。也可基于上网用户的其他资料如电话号码、电子邮箱地址、网络账号等完成随机抽样。

（3）被访者驱动抽样法，异于传统取样方式，组合滚雪球抽样与样本赋予权重的数学模型，以弥补非随机抽样缺陷的抽样方法。首先估计样本衔接的社会网络关系，再依据关系结构衍生受访者（称为“种子”）并激励其荐举定额的样本群体来填写E问卷。历经多次荐举后，“种子”衍生不再受制于样本偏差。相较传统抽样法，被访者驱动抽样在研究艾滋病患者、吸毒者等“隐秘群体”领域具有检索数据的高效性和精确性（余富强等，2019）。

（二）调查执行阶段

该阶段可划分为发布问卷、异常监测与调整、回收问卷、后续驱动监测四部分。调查者将定稿问卷投放给既定样本，即刻回收样本填答数据至云端（章国华，2008）。实时监测后台数据并核查，以便调整异常情况。全部回收过程告竣后，调查者可通过电话回访、短信回访等方式；结合自身常识与知识储备，剖断已回收问卷是否有重新调整的必要，完善问卷调查，保证数据质量（董贝蓓，2016）。

1. 发布问卷

传统纸质问卷调查受限于时间与空间，调研团队常以小组分工模式发布问卷，即划定样本群体后，调研团队针对子样本群体分派各调研小组开展问卷调查。该模式能够优化团队人员配置，提高工作效率，便于问卷数据管理。建立在互联网技术之上的网络问卷调查拥有超越时空的特性（李芳，2017），调研团队难以实现常规小

组分工。聚焦于此，项目组在实践探索中，可以实行“分群驱动”问卷发布策略，推动小组分工方式创新（袁德霞，2010）。

分群项目模式即创建多个问卷副本链接，同时将样本随机再分类产生子样本，以电子邮件、问卷链接、问卷二维码等形式发布给子样本，子样本填答相应问卷链接产生多个数据集（图2），初步分类调查数据。后续数据处理方式包括以下两种：（1）将数据集录入总问卷中，获取总数据，运行并完成数据分析；（2）汇总数据编码，导入统计分析软件，完成数据分析。在本次调研中，项目组成员分为四个小组分工发布问卷，实施分群驱动发布策略。将样本分为四组，复制原问卷链接，产生四个子问卷后，小组以电子邮件发送各自问卷链接，发布问卷调查。

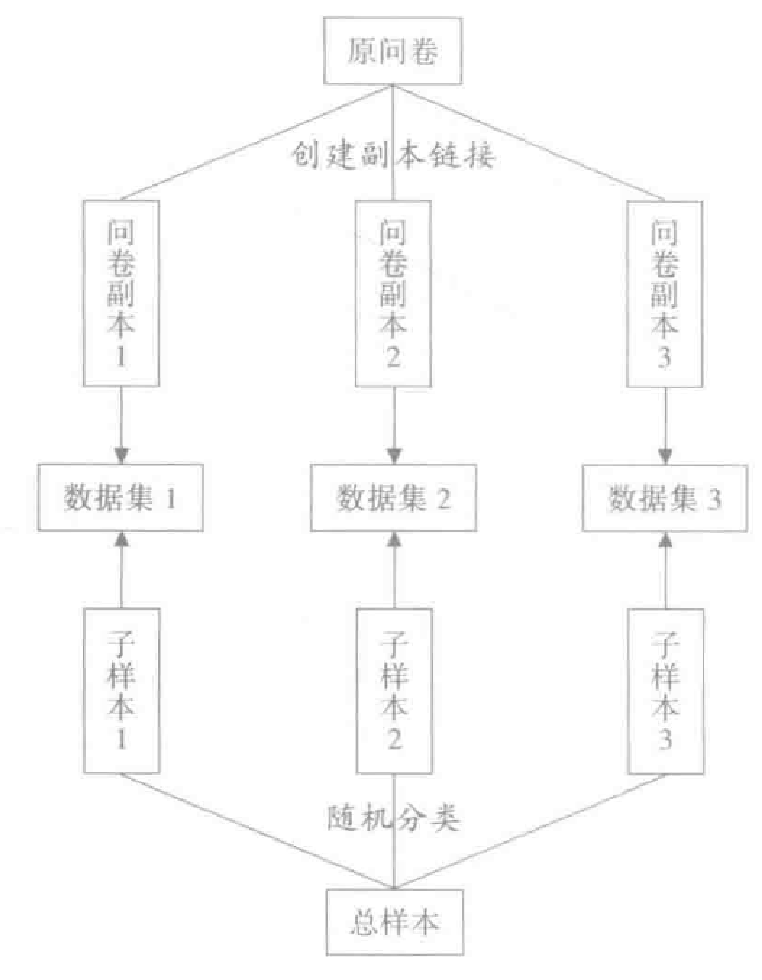


图2　分群项目的使用流程

使用分群项目有以下三个突出优势。(1) 便于管理问卷数据。不同于单一问卷链接-单一数据集的问卷发放模式，实施分组项目在调查执行阶段即可实现问卷数据预分类，减省数据分析步骤的时间成本，提高数据清洗-数据分析的效率。即时获得的分类数据集有助于项目组诊断数据异常，有针对性地反馈调整。(2) 高效实施回收驱动。有别于传统调查面对面或电话联系回收，E问卷的回收驱动涉及全网域，加之调查员与样本缺少接触，问卷填答率难以保证。适当的激励措施将有效解决这一难题。不同形式的激励手段，如货币（现金、支票）或非货币（礼物、积分点以及奖券等），有效性因人而异（罗玮、罗教讲，2015）。与无分别化激励措施不同，分群项目可实现基于样本特征的预分类，精准施行激励策略，发挥激励有效性。(3) 实现调查工作分组。在纸质问卷调查中，调查员依据区域或时间条件协商工作分组，能有效加快工作进度；而网络问卷调查跨时空属性，招致分组凭证缺失，在问卷发布-回收阶段不便于内部分工。分群项目将完整调查工作划分为子调查的集合，优化分工协作，提高效率效能。

2. 驱动监测

在网络调查中调查者无法有效调控问卷的派发和填写。在此过程中，调查对象出于个人意愿和动机趋向等原因，对调查重视度有异（郭继志等，2006），造成问卷填写质量良莠不齐。调查者可以调动填答激励和预设筛选两种手段，管控数据质量（方佳明，2013）。

在回收问卷阶段，恰当的激励手段不仅能增加网络问卷应答率，也为数据质量优化增添动力。伴随移动支付技术成熟，原有的现金奖励方式在网络问卷中成为可能。电子红包作为网络问卷现金

奖励的载体，直接将酬金存入被调查者账户，增加潜在受访群体的期望值（吴陈，2019）。设置定额报酬，能切实加强调查对象的参与度和问卷扩散意愿，同时也需防范利益驱使下恶意填答问卷的风险。实际操作过程中，须健全激励流向机制，让酬劳给予与问卷审核齐头并进。警惕问卷数量发生非常态爆发式增长，在苗头初现时，立即暂停接收问卷，追根溯源。根据起因调节、整固问卷扩散路径，以防大量无效或低质量问卷流入数据库。本次调查应用事后激励形式（郭继志，2006），提供小额电子现金红包。调查对象完成问卷并点击“提交”按钮后，系统自动跳转至领取红包页面。为避免重复填答，防控“刷单”行为（邵国松、谢珺，2021），保证问卷有效性，问卷设置同一IP地址或同一设备填答上限为1次，且当每刻钟内填写数量超过50份时，平台将暂停回收，以便项目组分析根由、排查隐患。

原始调查数据生成后，调查者应仔细检查回收的问卷，掌握其内外部特征，甄别、筛除失实数据。（1）网络问卷外部特征，即原始数据中存储的有关问卷填写基本状况的信息，如填写时长、地点等。问卷填写时长是判断数据异常的重要外部特征之一——适中的填写时间是问卷质量的有力保障，而填写时间异常的问卷数据的质量存疑（罗玮、罗教讲，2015）。问卷填答率也是评判异常的重要依据，调查者应刨除填答率极低、数据分析价值不高的问卷个案。若是样本问卷填写完整度普遍较低，在排除技术原因后，需重新考虑问卷是否存在措辞晦涩或表述不清的条目以及发放对象对此研究的接受程度，择取最佳抽样方式，激发填答积极性，提升原始数据质量水平（郭继志等，2006）。（2）网络问卷内部特征，即调查所围绕维度以及问卷设计中问题的跳转逻辑。目前大部分专业问卷平

台支持在问卷中插入陷阱题，检验所收样本有效性；大数据背景下，平台运用AI学习算法、人工智能参评量表，对已回收数据质量和置信度自动打分，同样能增加数据精确性（游培松等，2021；肖强、方国正，2021）。调查者可预先设置互斥性陷阱题，若作答出现前后逻辑失合或自相矛盾的情况，则视为未通过逻辑检验，问卷数据不予采纳。例如，案例所用问卷纳入的两个陷阱题项（图3）分别是“在大学期间未参与过兼职”与“在大学期间有过兼职受骗经历”，标记并删除对两个题项皆选择高认同度的问卷，保证样本数据信度。

图3　双题规则陷阱题

（三）分析写作阶段

此阶段囊括数据导出、数据清洗、数据分析、报告撰写。处理并存储完相关数据后，输出E问卷动态信息，导出调查数据。历经清洗筛选后将其录入统计软件，剖释数据，最终撰写调查报告，筑牢事业发展的坚实基础。网络调查所得资源能汇聚云端，以数据文件形态即时上传到数据库。然而此类数据缺乏数据集合与变量集合的格式规范，异于统计软件通用文件格式，无法直接运用其中，二者间存在数据对接问题（游培松等，2021）。

目前，部分专业网络问卷平台（如问卷星、腾讯问卷）提供问卷调查结果自动转化为常用统计软件数据格式的服务，推动软件间的无缝对接。不仅如此，近年来平台积极探索问卷调查、数据分析、报告撰写一体化机制，无需调用统计软件，在平台上就能完成常规数据分析（肖强、方国正，2021）。本次调查共回收数据502条，经预设筛选后保存有效数据498条，应用平台的相关分析、回归分析、方差分析等数据分析功能（图4），实施操作并根据导出的默认分析报告完成调研报告撰写。

通用方法		问卷研究		可视化	
频数	分类汇总	信度	效度	散点图	直方图
描述	交叉(卡方)	多选题	单选-多选	箱线图	词云
相关	线性回归	多选-单选	多选-多选	误差线图	P-P/Q-Q图
方差	t检验	项目分析	权重	Roc曲线	象限图
单样本t检验	配对t检验	验证性因子分析	对应分析	帕累托图	簇状图

图4　部分数据分析功能展示

当前国内应用广泛的专业问卷调研网站如问卷星、腾讯问卷及问卷网等，继承PC时代的基因，兼具编辑、发布、回收的能动性与数据处理方式的标准统一性等技术优势（Jolene et al.，2009），支持学术积淀相对缺失的企业和个人开展网络问卷调查，助力平台自身与网络调查的繁荣发展（孙秀林、施润华，2016）。以案例选用的问卷星平台为例，其致力于深挖数据价值，提供强大的数据收集、存储和分析工具，拥有数以千万计的用户体量。自2006年上线至今（截至2021年12月），用户累计发布了超1.43亿份问卷，累计回收超113.28亿份答卷（参见问卷星网站：https：//www.wjx.

cn/)。简易、高效的平台功能扩大了社会调查的参与圈，推动社会调查从传统社会科学活动向综合性学科活动演进。

处理脱离专业问卷调研平台的调查数据时，可采取传统编码、转码操作将数据导入统计软件。也可启用平台复制问卷功能，手动录入数据至问卷，充分发挥云平台智能转化、分析数据的效能，节省时间和人力成本（邵孟良，2014）。

四、唯实励新：E问卷实务发展展望

当今，共享化技术和网络化信息深深植根于社会土壤（卢锦运等，2015），为E问卷承办制度化和交互可视化发展擘画了蓝图。健全系统完备、有机衔接、长效管用的E问卷承办制度体系，构建多元、畅通、有序的交互可视化渠道，将拓宽网络问卷调查形式。多层次、宽领域地力促E问卷参与广度扩张，夯实E问卷实务规则框架，激发E问卷操作实务中的朝气活力。

其一，E问卷的承办制度化趋势。当前网络问卷平台正走向更加成熟更加定型的道路，倚仗问卷平台的技术支撑可初步实现E问卷实务操作的全盘流程。经过参与式观察发现，部分平台推出了代理机构协同调查功能（Stephen & Michael，2007），诸如问卷设计、抽取样本、数据分析、报告撰写等均可付费后委托平台方专业调查团队完成。E问卷的承办代理服务，将大幅降低网络问卷调查的技术门槛，搭建社会调查学科间交流的桥梁，扩大调查的社会开放度和参与度（邵孟良，2014）。但归咎于相关行业标准付之阙如，在商业利益驱策下，平台提供的代理服务难免会有为迎合目标用户口味而更动原调查内容、增加倾向性信息等信息异化现象存在。制度的缺失、标准的失效、平台的缺位是此现象发生的导火索。因此，

构筑高密度、上下联动、横向协作的监督检查网络迫在眉睫。只有在E问卷承办制度设计上下功夫、在敦促落实上出真招、在优化协同上用气力，才能将收集的成果转化为推动E问卷乘势而上的强大动能。

其二，E问卷的交互可视化趋势。随着移动通信技术的普及推广，高像素智能终端进入千家万户，带来稳定、自然、高质量的即时通讯效果，线上视频的交互式E问卷访问调查技术成本逐渐降低（Jarvis，2010）。伴随多种无线接入技术演进集成方案与移动端视频会议系统深入研究的同频共振，基于VR技术的元宇宙场景已初具雏形（张洪忠等，2021）。未来网络问卷调查可打破屏幕空间障碍的制约，将不同地方的人链接到同一个屏幕空间中，实现面对面式交互。这种同屏式互动、页面共享、场景批量管理功能，可以使调查者"一对一""一对多"地"手把手"协助调查对象进行问卷填写。对于问卷中的疑难问题也与线下调查趋同，可由调查员进一步展开说明，极大地突破了网络调查问卷真实性盲区。此外，随着面部表情识别技术的不断发展，AI可精准捕捉调查对象填写过程的表情动作进而实时校验填答结果的真实性。

移动互联、应用生态已点燃数字革命新引擎，大数据、区块链、数字孪生等技术日趋成熟，网络空间无远弗届。在未来网络问卷调查开展和推行过程中，应聆听这些新兴技术思维的召唤，回答时代之问。我们可借力全息互联网、虚拟现实、多感官交互，营造沉浸环境，为E问卷发展注入新能量；同时需强化承办制度意识，增强制度执行力，护航E问卷行稳致远。

参考文献

阿迪力·努尔，2012，《浅谈调查问卷设计中的有关技巧》，《统计科学与实践》第6期。

陈云松等，2016，《大数据机遇与宏观定量社会学的重启》，《贵州师范大学学报（社会科学版）》第6期。

董贝蓓、刘海燕、董哲，2016，《网络问卷调查有效性研究——基于大学生对土地利用信息心理需求的调查》，《山东理工大学学报（社会科学版）》第3期。

董海军、耿宇，2018，《移动互联网+问卷的应用特点与发展》，《晋阳学刊》第3期。

方佳明，2013，《网络调查中激励的有效性：基于荟萃分析的结果》，《管理评论》第2期。

方佳明、邵培基，2006，《一种评估网络调查适用度的方法》，《清华大学学报(自然科学版)》第1期。

郭继志、阎瑞雪、宋棠，2006，《网络调查方法的优势与局限》，《中国社会医学杂志》第1期。

何哲，2020，《网络治理在公共突发危机事件中的作用》，《中国广播》第9期。

蒋笃君，2021，《红色基因：百年大党的“根”与“魂”》，《人民论坛·学术前沿》第19期。

李芳，2017，《大数据时代传统统计抽样调查面临的机遇和挑战》，《财经界》第33期。

李锐、宋铁英，2005，《国内网络调查研究分析》，《情报科学》第6期。

李巍，2009，《网络调查研究方法应用效果的实证研究》，山西师范大学硕士学位论文。

刘慧园，2019，《积极培育网络治理中的四个基本认同》，《中国党政干部论坛》第12期。

刘洋滔，2021，《基于移动互联网的多元社会调查探索——以四川省级绩效评

价为例》,《财政监督》第18期。

刘玉拴, 2019,《网络文化安全治理体系研究》,《中共中央党校》第186页。

卢锦运、何怀金、刘明, 2015,《支持逻辑跳转及填答约束的网络问卷调查系统设计与实现》,《软件导刊》第14期。

卢锦运、何怀金、刘明, 2015,《支持逻辑跳转及填答约束的网络问卷调查系统设计与实现》,《软件导刊》第1期。

罗玮、罗教讲, 2015,《新计算社会学:大数据时代的社会学研究》,《社会学研究》第3期。

吕正娟, 2012,《网络问卷调查的规划、设计与实施》,山东师范大学硕士学位论文。

邵国松、谢珺, 2021,《我国网络问卷调查发展现状与问题》,《湖南大学学报(社会科学版)》第4期。

邵孟良, 2014,《大负载下的问卷调查系统的性能优化》,《电脑知识与技术》第3期。

孙秀林、施润华, 2016,《社会学应该拥抱大数据》,《新视野》第3期。

王齐齐, 2021,《论网络社会视阈下法的价值》,《重庆理工大学学报(社会科学)》第8期。

王莹、万舒晨, 2016,《大数据时代抽样调查面临的挑战与机遇》,《统计与信息论坛》第6期。

卫海英、杨智勤, 2001,《网上调查误差的成因及减少措施探析》,《中国统计》第7期。

吴陈, 2019,《网络匿名性对消费者非伦理行为形成的影响机制研究》,武汉理工大学硕士学位论文。

肖强、方国正, 2021,《关于视频会议系统跨系统互联的设计与实现》,《数字通信世界》第11期。

闫建、郑迦元, 2020,《"互联网+"背景下的地方政府治理现状:问卷调查与统计分析》,《重庆行政》第3期。

游培松等, 2021,《VoLTE高清视频通话在鼠疫应急指挥中的应用》,《医学动物防制》第10期。

余富强、胡鹏辉、杜沙沙, 2019,《网络问卷调查的数据质量控制研究》,《统计与决策》第16期。

袁德霞, 2010,《浅谈互联网调查》,《网络财富》第4期。

张洪忠、斗维红、任吴炯，2022，《元宇宙：具身传播的场景想象》，《新闻界》第1期。

张静，2009，《谈统计调查的新方法——网络调查及其应用》，《辽宁师专学报（社会科学版）》第6期。

张青、瞿莉，2020，《网络学习社区的构成要素探索——以教师研修社区为例》，《教育信息技术》第12期。

章国华，2008，《手机调查的技术初探》，《统计研究》第11期。

Bosnjak, Michael & T. L. Tuten 2003, "Prepaid and Promised Incentives in Web Surveys." *Social Science Computer Review* 21(2).

Chisnall, P. M. 2007, "Mail and Internet Surveys:The Tailored Design Method." *Journal of Advertising Research* 47(2).

Jarvis, M. 2009, "Encyclopedia of Survey Research Methods" In Paul J. Lavrakas (ed.), *Encyclopedia of Survey Research Methods*. Thousand Oaks.

Jolene, D. S. et al. 2009, "Open-Ended Questions in Web Surveys: Can Increasing the Size of Answer Boxes and Providing Extra Verbal Instructions Improve Response Quality?" *The Public Opinion Quarterly* 73(2).

Katja, L. M. et al. 2008, "Web Surveys versus other Survey Modes: A Meta-Analysis Comparing Response Rates." *International Journal of Market Research* 50(1).

Mick, P. C. 1997, "Survey Introductions and Data Quality." *The Public Opinion Quarterly* 61(2).

Stephen, R. P. & E. W. Michael 2007, "Mixed-Mode Contacts in Web Surveys: Paper Is Not Necessarily Better." *The Public Opinion Quarterly* 71(4).

Tse-Hua, S. & F. Xitao 2008, "Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis." *Field Methods* 20(3).

Zhifeng, G.、A. H. Lisa & X. Jing 2016, "Online Survey Data Quality and Its Implication for Willingness-to-Pay: A Cross-Country Comparison." *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* 64(2).

王金水，吴俞晓，许琪．年龄-时期-世代模型的发展历程与社会科学应用[M/OL]//
赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

年龄-时期-世代模型的发展历程与社会科学应用

王金水　吴愈晓　许琪[①]

摘要：作为社会变迁研究的有力分析工具，年龄-时期-世代（APC）模型近年来得到了国内社会科学学者的广泛关注。由于这三个时间维度存在完全共线性关系（年龄=时期-世代），求解该模型始终面临严峻挑战。本文系统回顾了传统APC模型的基本原理和五种最常用的求解方法：受限广义线性回归法（CGLM）、代理变量法、内生因子法（IE）、分层交叉随机效应APC模型（HAPC）以及最新的APC交互模型（APC-I）。此外，本文总结了国内社会科学研究中APC模型的应用情况，并针对国内数据来源的特点介绍了不等间距数据在APC模型中的应用。最后，本文分析了未来APC模型可能的发展前景以及在国内社会科学应用中的机遇和挑战。

关键词：年龄-时期-世代模型；社会变迁；世代效应；不可识别问题

①作者简介：王金水，南京大学社会学院博士研究生，联系方式：njuwjs@smail.nju.edu.cn；吴愈晓，南京大学社会学院教授，联系方式：yxwu2013@nju.edu.cn；许琪，南京大学社会学院副教授，联系方式：xuqi@nju.edu.cn。

Abstract: Age-period-cohort model (APC) has been widely adopted by domestic social science research in recent years. Although APC model has a long history in western literatures, there is little agreement on how to adequately analyze APC model. This article begins with an overview of classical APC model, discussing the identification problem. We then outline five major techniques to solve the identification problem: (1) restricted generalized linear regression (CGLM); (2) proxy variable method; (3) intrinsic estimator (IE); (4) hierarchical cross-random effect APC model (HAPC); (5) APC interaction model (APC-I). In addition, this article summarizes the Chinese social sciences literature that used APC model, and introduces the application of unequal spacing data in APC model according to the characteristics of Chinese data sources. Finally, we analyzes the possible development prospects of the APC model in the future, as well as the opportunities and challenges in the application of Chinese social sciences.

Key words: Age-Period-Cohort Model; Social Change; Cohort Effect; Identification Problem

年龄-时期-世代模型（Age-Period-Cohort model，以下简称APC模型），也被称为年龄-时期-队列模型，是研究社会变迁的有力分析工具。区别于一般的社会变迁的研究[①]，年龄、时期和世代分别反映了变迁（时间）的三个维度，其中年龄效应（age effect）是有关个体生命历程改变以及与年龄相关的变迁效应，即特定因变量的年龄差异；时期效应（period effect）是指在特定调查年份的宏观社会、经济、文化状况对所有被调查者产生相似的影响，即特定

①通常情况下对于社会变迁的研究只会关注到一个维度，即时期的维度。

因变量在不同调查年份的差异；世代效应（cohort effect）是指早年生活条件或社会经历对在同一年份（时段）出生的人产生的影响，即特定因变量在不同出生世代的差异（Ryder，1965；Yang，2008）。

弗洛斯特（Frost，1939）最早进行了年龄、时期和世代分析，随后这一重要的分析视角被广泛应用于流行病学的研究之中，在20世纪50年代后开始应用于人口学和社会学的研究之中（Fosse & Winship，2019）。在国内的研究之中，据苏晶晶和彭非（2014）的梳理[①]，沈其君（1992）最早在公共卫生与流行病学研究领域介绍了APC模型。迄今为止，国内学者对于APC模型的使用也主要集中于公共卫生与流行病学的研究，在社会学等社会科学研究中的应用相对较少。这主要有两方面的原因。其一是数据的限制，APC模型的主要作用是描述变迁的趋势，这就要求有连续多期的数据，而我国的社会调查和数据共享工作起步相对较晚，早期难以提供充足的数据供使用。但是，近年来诸如中国综合社会调查（CGSS）、中国综合社会状况调查（CSS）等大型国内调查已经持续多年，为国内学者提供了比较丰富的数据，这也为使用APC模型提供了一个重要契机。可以观察到，最近几年来国内社会科学界对于APC模型的使用明显增多。其二是APC模型自身的限制，由于年龄、时期和世代三者之间存在完全共线性（年龄=时期-世代），所以会导致模型难以估计出唯一的解，也即存在“不可识别”的问题。虽然学者们已经提出诸多方法来解决这一问题，但是尚未形成共识，彼此的批评也从未间断（详见：Fosse & Winship，2019；Luo，2013等），这

①原文指出“余国培（1992）和沈其君（1992）最先在国内引入APC模型”但原文未列出相应参考文献，笔者也没有查找到余国培（1992）这篇文章。

也就导致很多学者还尚未接触APC模型就产生其不可靠的刻板印象，影响了APC模型在国内的应用。

为此，本文希望秉持中立的立场详细地阐述APC模型的原理及其应用，旨在让读者能够全面地认识到APC模型的价值和局限。接下来，本文第一部分将简要阐述为什么要使用APC模型；第二部分将回顾APC模型的基本原理，以及常见的APC模型估计方法及其最新进展；第三部分将梳理国内社会科学研究中APC模型的使用情况；第四部分将针对中国社会调查数据的特点，介绍不等间距数据在APC模型中的应用；最后一部分将对APC模型在国内社会科学中的应用进行展望。

一、为什么使用APC模型

（一）社会变迁多个维度的混杂效应

社会变迁是社会科学研究的重要议题之一，尤其是在中国体制改革以及经济飞速发展的背景之下，社会变迁研究更加具有理论和现实意义。然而，在很多关于社会变迁的研究之中，学者们只关注到了历时性的变化，也即前文指出的时期效应。这种研究取向忽略了社会变迁在时间维度的复杂性。社会变迁在时间维度上包含了年龄、时期和世代三种相互“混杂”（confounding）的效应，也即因为三者之间的特殊关系，会产生彼此制约效果。例如，我们观测到的年龄效应很有可能是世代效应（O'Brien，2011），而时期效应又会受到人口更替的影响，即新的出生世代不断接替老的出生世代（Ryder，1965）。但由于年龄、时期和世代三者之间的关系过于特

殊，尤其是年龄效应和世代效应，对于初次接触APC模型的读者可能很难区分，这里我们通过杨洋和兰德（Yang & Land，2013，pp. 9-10）的一个例子来进一步讲解三者的混杂效应。表1是年龄和时期的交叉表（也被称为列克西斯表），其中第一列a60、a70、a80、a90分别代表年龄60岁、70岁、80岁和90岁，第一行p1980、p1990、p2000、p2010分别代表时期1980年、1990年、2000年和2010年，表中的其他元素c1920、c1930等代表出生世代。例如年龄为60岁在1980年对应的出生世代为1920，在1990年对应的出生世代为1930。从表1中我们可以直观地观察到，在同一个时期，年龄效应和世代效应是很难区分的，类似的在同一年龄上，时期效应和世代效应也是很难区分的。例如，在1980年，60至90岁年龄的变化与1920至1890年世代的变化无法区分；在60岁组，1980—2010的时期变化与1920—1950的世代变化也无法区分。

表1　年龄-时期交叉表

	p1980	p1990	p2000	p2010
a60	c1920	c1930	c1940	c1950
a70	c1910	c1920	c1930	c1940
a80	c1900	c1910	c1920	c1930
a90	c1980	c1900	c1910	c1920

因此，研究者要想综合全面地理解社会变迁，单独考察一个维度的变迁效应是不准确的，需要在一个分析框架中同时考察这三种效应，方能更准确地了解某个特定时间维度的变迁情况以及各自的解释机制。而APC模型正是在同一框架下分析三种效应的统计方法。

此外，需要说明的是，虽然APC模型能够厘清社会变迁（时

间）的三个维度：年龄、时期和世代，但是“时间变量”本身包含着十分丰富的信息。换言之，年龄、时期和世代作为变量纳入模型并不意味着它们是真正的解释性变量，真正有解释力的可能是年龄背后的生物化老化（比如器官功能衰退）和社会性成熟（比如人生阅历丰富）、时期背后的重要社会性事件（比如2008年金融危机、2020年的新冠疫情）、世代背后成长的共同环境（比如成长于计划经济时代、成长于互联网时代等）。因而，APC模型更重要的目的是清晰地描述出特定因变量是如何随年龄、时期和世代变化的。

（二）世代效应的理论和现实意义

在社会科学研究之中，相比于年龄效应和时期效应，世代效应往往会被忽视。世代的英文是cohort，在人口学中，它常被译为“队列”，如出生队列指的是特定时期（某一年或某一段时间）出生的人口。其他学科的研究者可能会对队列一词略感陌生。Cohort一词的英文含义为“a group of people who share a common feature or aspect of behavior”，即“有共同特点或举止类同的一群人”，也可以被译作“世代”“同期群”和“同龄组”等。而对于“世代”一词大家可能会更加熟悉，有关“世代”或“代”问题的研究一直是社会科学所关注的重点内容。

20世纪50年代，社会学家卡尔·曼海姆（2002，pp.65—108）就提出了“代”（generation）的概念，这一概念与“世代”或“队列”具有相同的含义。曼海姆对代问题进行了详细的论述，他认为同一世代的个体因为早期的特定经历因而具有相似性；不同世代会因各自年轻时受到不同历史进程影响而产生思维和行动倾向的分化。根据曼海姆的理论，相同世代的个体因为童年时或早年社会化

过程中经历相同的社会情境，从而形成了内部一致的独有的文化或价值观意义上的“世代特征”，且具有顽强的持久性；而不同的世代之间，其特征则会呈现系统性的差异。而社会变迁则体现为“新的参与者在文化过程中不断出现”和“老的参与者从文化过程中不断消失”的过程（卡尔·曼海姆，2002）。

受到曼海姆的影响，美国社会学家莱德（Ryder，1965）也指出，社会结构和文化的变迁不仅仅是因时间推移而发生的变化，而且是一个社会的人口继替（succession）过程，即新的出生世代不断接替老的出生世代的过程。而广为流传的经典著作《大萧条的孩子们》就是分析了在1920—1921年这一世代中出生的个体，在历经美国大萧条的影响下一生的成长与发展（埃尔德，2002）。在此基础上也逐渐发展出生命历程的分析范式，而世代的视角则是生命历程分析范式中的重要组成部分（Giele & Elder，1998；Glenn，2005）。APC模型也恰恰是进行世代分析的重要工具。有关“世代”的问题或是说世代的问题，不仅是学术研究中的重点，在我们日常话语中也能体现出对它的重视。例如，我们日常话语中经常提及的“80后”“90后”现象本质上就是世代效应。

二、APC模型与求解方法

（一）经典APC模型及其参数识别问题

虽然弗洛斯特（Frost，1939）最早进行了APC的分析，他的研究也被视为APC研究的经典文章，但是在他的研究中只是通过列克西斯表（年龄和时期的交叉表）进行了描述性分析。而经典APC模

型（classical APC model，简称C-APC），也被称为多重分类APC模型（multiple classification model）或多重解释APC模型（accounting model），最早是由梅森等学者（Mason et al.，1973；Mason & Smith，1985）提出的，并且后续所有的APC分析方法都是在此基础上展开。

为了更好地理解传统APC模型，我们需要首先了解模型的数据要求以及数据结构。首先，APC模型需要有多期数据，而至少需要有多少期数据，现有的研究并没有给出明确的答案。其次，在数据类型上，统计汇总数据、截面调查数据以及追踪数据等均可应用于APC模型当中（Yang Y & Land，2013）。在数据结构上，我们可以通过年龄-时期交叉表来进行理解，如表2所示。该表中第一行代表时期，第一列代表年龄，每条对角线代表一个世代，可以证明，如果年龄有I个类别，时期有J个类别，那么世代将包含$K=I+J-1$个类别。表2展示的是一个$I=5$，$J=3$情形，相应地，表中共有7（$K=5+3-1$）个世代。

表2　年龄-时期交叉表

	$j=1$	$j=2$	$j=3$
$i=1$	$k=5$	$k=6$	$k=7$
$i=2$	$k=4$	$k=5$	$k=6$
$i=3$	$k=3$	$k=4$	$k=5$
$i=4$	$k=2$	$k=3$	$k=4$
$i=5$	$k=1$	$k=2$	$k=3$

如果我们用i作为年龄的下标，j作为时期的下标，k作为世代的下标，那么对于一个感兴趣的因变量Y_{ijk}，它与年龄、时期和世代之间的关系可以通过以下方程来表示：

$$Y_{ijk} = \mu + Age_i + Period_j + Cohort_k + \varepsilon_{ijk} \quad (1)$$

方程（1）中，μ为模型截距，Age_i代表i个年龄效应，$Period_j$代表j个时期效应，$Cohort_k$代表k个世代效应，ε_{ijk}为误差项。上述方程也将年龄、时期和世代均以分类变量的方式纳入模型，在具体分析时，可以参照回归分析的通常做法，将$I-1$个年龄虚拟变量、$J-1$个时期虚拟变量和$K-1$个世代虚拟变量纳入模型；也可参照方差分析的通常做法，令$\sum_{i=1}^{I} Age_i=\sum_{j=1}^{J} Period_j=\sum_{k=1}^{K} Cohort_k=0$。可以证明，虚拟变量编码和方差编码是等价的，但在有关APC模型的文献中，方差编码更常见，因此，我们在下文统一使用方差编码进行分析。

不过，正如梅森等学者指出的，即便我们按照虚拟变量编码或方差编码的通常做法对模型参数进行常规限定，该模型也是不可识别的（Mason et al.，1973）。原因在于该模型包含$I+J+K-3$个待估参数，但仅包含$I+J+K-4$个不完全线性相关的自变量。由于年龄=时期-世代，该模型不满秩（not full rank），常规统计方法无法给出所有参数的唯一解；或者说存在无穷多个解，这些解可以无差别地拟合观测数据。在没有附加信息的情况下，模型无法判断哪个解是最优解，这就是APC模型的参数“不可识别”的难题。

（二）APC模型求解方法

虽然经典APC模型存在参数识别问题，但由于对年龄、时期和世代效应进行分析具有独特的魅力，这激励了众多学者对这一问题进行探索，因而APC模型的发展历程可以被认为是不断探索如何解决“不可识别”难题的过程。本文接下来将介绍五种主要的应对策略，其中CGLM方法、代表变量方法和IE的方法在本质上没有改变

传统APC模型的设定，只是在求解过程中进行理论限定或数学限定；HAPC和APC-I更改了传统APC模型的设定，因而我们将前三种应对策略称为“方法”，而后两种称之为“模型”。

1. CGLM方法

CGLM方法，也称受限广义线性回归法（constrained generalized linear regression）。APC模型无法求解是因为模型的秩仅比满秩少1，如果对模型参数增加一个额外的限制条件即可求解（Mason et al.，1973）。因此，CGLM方法的主要原理就是对待估计参数施加一个限制条件，也即“受限”，并且该方法不仅适用于因变量是连续变量的线性回归模型，也适用于因变量是二分变量、计数变量或其他变量类型的广义线性模型。对模型参数施加某种限制条件，一个最简单的做法就是将两个参数设为相等，如设置$Age_1 = Age_2$、$Period_1 = Period_2$或$Cohort_1 = Cohort_2$。

但CGLM方法最大的问题就是如何找到这种限制条件，也即我们为什么可以认为模型中某两个参数是相等的？这必须要基于扎实的理论，例如在我们分析某一问题时有明确的、可靠的理论告诉我们模型中60岁年龄组和70岁年龄组的效应完全一致，或是2000年和2005年的时期效应完全一致，再或是1980年和1985年的世代效应完全一致。但是，在实际研究中我们很难找到这样的理论限定。因而CGLM方法的应用也就相对较少。不仅如此，一些模拟研究发现，CGLM法对限制条件非常敏感，将两个存在细微差异的参数设为相等也会导致非常有误导性的结果。正因为这一明显的缺陷，所以该方法在实际应用中的争议很大（Glenn，1976；Rodgers，1982）。

2. 代理变量法

当学者们意识到了CGLM法的缺陷后，又提出了新的求解方法——代理变量法（Heckman & Robb，1985；Kahn & Mason，1987）或奥布莱恩（O'Brien，2011）所说的特征变量法。该方法的基本思路是寻找一个年龄、时期和世代以外的变量，来代替年龄、时期或者世代，如果该变量能完全代表年龄、时期或世代中的一个对因变量的影响，且与另外两个变量不存在完全共线性，那么就可以用它替换被代理的变量求解APC模型。例如，如果出生世代的规模是导致世代对因变量产生影响的唯一因素，且该变量与年龄和时期不存在完全共线性，那么就可以用它替换世代进行APC分析。可以看出，该方法与CGLM方法相似，不是从数学或统计学原理的角度来解决求解的问题，而是依靠理论来解决不可识别问题，但是在实际应用中，很难找到符合要求的代理变量。

为了提高代理变量法的可行性，温希普和哈丁提出了一种基于机制分析的APC模型求解方法（Winship & Harding，2008）。这一方法的思路是首先基于理论找到一组年龄、时期和世代对因变量产生影响的中介变量，如果这组中介变量能够穷尽年龄、时期和世代中的一个变量对因变量的影响机制，就可以获得APC模型的解。但该方法依然要求研究者能够穷尽年龄、时期和世代中的一个变量对因变量的影响机制，这在实际应用中往往也很难实现。正因如此，该方法虽有理论意义，但在实践中的应用很少。

3. IE方法

以上两种方法都要研究者做很强的假定，在实际应用中往往很

难实现，这也就大大限制了它们的应用前景。为了寻找新的求解思路，在2000年以后，西方学术界兴起了一股使用统计方法求解APC模型的风潮。屈佩尔等学者（Kupper et al.，1985）从年龄、时期和世代所组成的设计矩阵角度证明了APC模型可能存在确定的解。正是在此基础之上，杨洋等学者（Yang et al.，2004）提出了IE方法，又称内生因子法（intrinsic estimator），其本质上是一种主成分求解法，求解过程共分五步。

第一步：对模型自变量进行主成分分析。通常来说，对n个不存在完全共线性的变量进行主成分分析可以得到n个主成分，每个主成分对应一个非零的特征值，但APC模型包含的$I+J+K-3$个自变量中有一个是冗余的，因此，对它们进行主成分分析只能得到$I+J+K-4$个非零特征值，与之对应的有$I+J+K-4$个独立的主成分，另一个与零特征值对应的主成分可以由其余$I+J+K-4$个主成分线性表示出来。

第二步：根据主成分分析的结果，把$I+J+K-3$个自变量构成的向量空间分为两个部分，一是与$I+J+K-4$个非零特征值对应的非零向量空间（nonnull subspace），二是与零特征值对应的零向量空间（null subspace）。

第三步：以非零向量空间中的$I+J+K-4$个特征向量为自变量拟合回归模型，这样可以得到$I+J+K-4$个回归系数。

第四步：将与零向量空间对应的系数设为零。

第五步：将第三步得到的$I+J+K-4$个回归系数与第四步人为添加的一个回归系数零合并起来，然后通过矩阵旋转得到与原APC模型中的$I+J+K-3$个自变量对应的所有解。

在上面五个步骤中，第四步将与零向量空间对应的系数设为零是

求解过程的关键。杨洋等人（Yang & Land，2008）的解释是，零向量空间对因变量没有影响，因此将与该向量空间对应的系数设为零。但罗丽莹（Luo，2013a）的研究却发现，这一设定实际上增加了一个非常武断的假定，这个假定不仅很难在实践中得到满足，而且会随着年龄和时期类别数量的变化而变化，因此IE的求解结果通常是有偏的。后续的研究还发现，IE是在所有可能的解中选择了一个参数平方和最小的解，如果用图形展示，这个解在所有可能的解中距离向量空间原点的欧氏距离最短，但距离最短与最优是两个不同的概念（Fosse & Winship，2019）。而且，向量空间中原点的位置取决于自变量的编码方法，例如方差编码与虚拟变量编码的原点就不一样，而且同样是虚拟变量编码，以哪个组作为参照组也会改变原点位置，这导致在不同编码方案下使用IE会得到完全不同的结论（Luo，2013a）。简言之，IE方法在本质上没有改变传统APC模型设定，即依旧是方程（1）的模型设定，只是在运算原理上进行了调整。

4. HAPC模型

HAPC模型，又称分层交叉随机效应APC模型（hierarchical crossed random-effect APC model），该方法是由杨洋和兰德（Yang & Land，2006，2008）提出的。其主要原理就是利用多层次模型，将年龄、时期和世代三者中的一个或两个放到第二层，从而使得三者不在同一测量层次，这样就避开完全共线性问题。在具体实践中，杨洋和兰德建议将年龄作为个体层次的变量，即多层模型的第一层，将时期和世代作为更高层次的变量，即多层模型的第二层，来拟合分层交叉随机效应。这一方法改变了传统APC模型的设定，其数学表达如下：

第一层模型：

$$Y_{ijk} = \beta_{0jk} + \beta_1 Age_i + \varepsilon_{ijk} \tag{2}$$

其中β_{0jk}为模型截距，Age_i代表i个年龄效应，ε_{ijk}为误差项。研究者可以根据需要在这一层控制需要控制的其他个体层面的自变量。

第二层模型：

$$\beta_{0jk} = \gamma_0 + Period_{0j} + Cohort_{0k} \tag{3}$$

其中γ_0为模型截距，$Period_{0j}$代表第j个时期效应，且$Period_{0j}$服从正态分布，$Cohort_{0k}$代表第k个世代效应，且$Cohort_{0k}$服从正态分布。将上述式（3）带入式（2）就可以得到HAPC模型的最终方程（5）：

$$Y_{ijk} = \gamma_0 + Period_{0j} + Cohort_{0k} + \beta_1 Age_{ijk} + \varepsilon_{ijk} \tag{5}$$

这一方法自提出以后得到了非常广泛的应用。我们关注到，大多数国内学者都在使用该方法求解APC模型（参见本文第三部分）。这或许是因为多层线性模型已经是一种非常成熟的统计方法，且将时期和世代作为更高层次的变量也比较符合人们的主观感受。

但是，也有学者针对HAPC方法提出了质疑。贝尔和琼斯（Bell & Jones，2014，2015）通过数据模拟研究发现，在世代效应呈线性变化或近似线性变化的情况下，使用HAPC通常会得到无世代效应的结论，即便在世代效应呈现出明显的非线性变化的情况下，使用HAPC也很容易抹除这种变化背后的线性趋势。他们认为，导致这一结果的主要原因是，多层线性模型在求解时需要最小化误差方差，随机效应作为误差方差的一部分必须服从这一原则。在APC模型中，世代的类别数量往往大于年龄和时期（$K = I + J - 1$），因而，将世代效应中的线性趋势归入时期或年龄的名下总能减少误差方差，这导致使用HAPC很容易得到世代效应不存在或没有

明显变动趋势的结论。罗丽莹和霍奇斯（Luo & Hodges，2020a）也指出HAPC模型的缺陷，他们认为HAPC之所以能得到APC模型的唯一解，是因为它在求解过程中已经对模型参数施加了某种隐含假定，因此，这种方法与CGLM并无本质区别，只不过HAPC不需要研究者人为添加假定，因而产生了一种不需要假定即可求解的错觉。他们证明，如果将年龄、时期和世代中的一个设为随机效应，等价于假定该随机效应的线性趋势为零，或者说假定该效应不会随时间呈现出明显的上升或下降趋势。如果将年龄、时期和世代中的两个设为随机效应，模型会根据随机效应的非线性部分的求解结果将两个随机效应中的一个的线性趋势设置为零，如果按照杨洋和兰德的建议，将时期和世代设为随机效应，那么在大多数情况下，研究者会得到线性趋势为零的世代效应。

总之，HAPC模型虽然得到了非常广泛的应用，并且模型背后的原理也看似相对合理，但也有不少质疑的声音。尤其是当世代效应为线性趋势时HAPC的方法可能会存在较大的偏误，因而也提示研究者在使用时要慎重考虑。

5. APC-I模型

APC-I模型（age-period-cohort interaction model），也称APC交互模型，是由罗丽莹和霍奇斯（Luo & Hodges，2020b）提出的最新的求解APC模型的方法。这一求解方法也更改了传统APC模型的设定，该模型使用年龄效应和时期效应的交互项替换了原始的世代效应，这也是它被称作APC交互模型的主要原因。模型的数学表达如下：

$$Y_{ijk} = \mu + Age_i + Period_j + Age_iPeriod_j + \varepsilon_{ijk} \tag{6}$$

方程（6）中，μ为模型截距，Age_i代表i个年龄效应，$Period_j$代表j

个时期效应，$Age_iPeriod_j$代表年龄和时期的交互效应，ε_{ijk}为误差项。该模型最大的特点是模型设定不再包含世代效应，而是利用年龄和时期的交互效应来替代世代效应，因而完美避免了共线性问题。但大家可能存在疑问：为什么年龄和时期的交互效应就可以代表世代效应？具体来说，根据APC模型的数据结构（表1），世代是由年龄和时期的交互分类决定的，据此可以将$I \times J$个年龄与时期的交互项$Age_iPeriod_j$分解到对应的世代（表3），然后对属于同一个世代的交互效应求均值即可算出世代效应（例如，对于表3中的世代3，对$Age_3Period_1$、$Age_4Period_2$和$Age_5Period_3$求均值即可得到世代效应$Cohort_3$的估计值）。不仅如此，根据表3，研究者可以得到不同时期的世代效应，因而可以分析世代效应随时期的变动趋势。

表3　APC-I模型中年龄和时期的交互项与世代效应的对应关系

	j=1	j=2	j=3
i=1	$Age_1Period_1, k$=5	$Age_1Period_2, k$=6	$Age_1Period_3, k$=7
i=2	$Age_2Period_1, k$=4	$Age_2Period_2, k$=5	$Age_2Period_3, k$=6
i=3	$Age_3Period_1, k$=3	$Age_3Period_2, k$=4	$Age_3Period_3, k$=5
i=4	$Age_4Period_1, k$=2	$Age_4Period_2, k$=3	$Age_4Period_3, k$=4
i=5	$Age_5Period_1, k$=1	$Age_5Period_2, k$=2	$Age_5Period_3, k$=3

罗丽莹和霍奇斯（Luo & Hodges，2020b）认为，他们提出的新APC模型及求解方法很好地体现了莱德的世代理论含义。根据这一理论，世代效应缘于年龄与时期之间的交互，且会随时期发生变化（Ryder，1965）。除此之外，APC-I模型还有一个其他APC模型所不具备的优势。对于每一个出生世代，其他的APC模型只能得出一个世代效应的估计系数，因此，它们只能假定每一出生世代的效

应在其整个生命历程中是恒定的。而APC-I模型可以在每一世代所处生命历程的不同阶段得到多个不同的世代效应估计值，因此，它使得研究者能够观察到世代效应随生命历程的动态变化过程，这为我们加深对于世代效应形成原因的理解提供了更丰富的信息。

由于APC-I是2020年提出的新方法，目前还没有关于该方法的进一步研究以及应用。因而也较少有质疑该模型的文献，但是这里还是要提醒大家，没有质疑的文献并不意味着这就一定是一个完美的方法，只是因为该方法提出得较晚。

（三）小结：求解APC模型是数学问题还是理论问题？①

综上，本文详细地回顾了APC模型常见的五种求解方法，这些方法的本质就是从不同的角度来应对参数的识别问题。APC模型的参数识别问题从根本上来看就是因为年龄、时期和世代效应之间存在共线性关系，为此只需要施加一个假定即可求解。这看似是一个数学问题，并且后来的学者们对于APC模型的讨论从形式上看也是数学（表达公式）或算法上的改进，但是这背后可能更多的是理论问题。因为，这些数学的改进或“算法”背后实际上都利用了有关年龄、时期和世代三者效应中的某一个或几个的理论性认识。这正如谢宇（Xie Y，1996）在评论查尔斯的《社会科学中的识别问题》一书中所指出的那样：统计学没有免费的信息，要么收集它，要么假定它。从上述的回顾可以看出，所有的APC模型的求解方法，都是在有限的信息的前提下试施加了不同的假定。只是有的明确表达了这种假定，而有的并没有明确表达这种假定而是内在隐含着。

因而APC模型的参数识别问题不仅是数学问题，更是理论问

①本部分内容得到了审稿专家的有益启发和建议，在此表示感谢！文责自负。

题。但这里所说的“理论”和我们一般在做研究中使用的社会学理论以及推导研究假设的理论不同，是指针对模型求解的理论。对于这一点，我们用大家更为熟悉的因果推断的方法做类比，例如：研究者在使用工具变量时，必须说明选择工具变量的理由，而工具变量的选择从统计上很难论证合理性，只能依靠理论来说明其是合适的，来论证工具变量只通过自变量一个途径影响因变量这一个关键假定。本文认为应用APC研究与之类似，在模型求解时，研究者需要额外的理论去论证所使用的求解方法是正确的。这就要求研究者要明确每一种方法背后的假定，表4具体列出了每一种方法背后的假定。在这5种方法之中，CGLM方法和代理变量法的假定属于言明的假定，也即研究者如果利用这两种方法必须先说明能够满足假定的原因。而IE方法和HAPC模型，他们的假定是暗含的假定，暗含于数理推导之中，对于使用者而言相对难以证明，尤其是IE方法。相比之下，APC-I模型的假定较弱，用年龄和时期的交互效应来替代世代效应也是莱德（1965）在提出世代效应时的重要观点。

表4　APC模型求解方法背后的假定

	假定
CGLM方法	在年龄、时期或世代变量中有两个变量值的效应是相等的，如：$Age_1 = Age_2$、$Period_1 = Period_2$或$Cohort_1 = Cohort_2$
代理变量法	其他变量能够代替年龄、时期或世代变量，或是说年龄、时期或世代变量的效应完全来源于其他变量
IE方法	将与零向量空间对应的系数设为零
HAPC模型	将年龄、时期和世代中的一个设为随机效应，或是说该效应不会随时间呈现出明显的上升或下降趋势
APC-I模型	世代效应是年龄和时期变量的交互效应

三、APC模型在国内社会科学实证研究中的应用

APC模型在国内更多地应用于公共卫生与流行病学或卫生经济学的研究之中，在社会学、人口学、传播学等社会科学中的应用相对较少。我们检索了特定范围[1]的使用APC模型的社会科学研究文献，如表5所示。从中可以发现，国内社会科学领域最早使用APC模型是在2007年，此后7年中没有研究使用该模型，2014年有两篇文章使用该模型，自2017年开始，使用APC模型的文章数量开始增多，这也得益于我国各种社会调查数据的积累。

此外，在学科与研究主题方面[2]，主要集中于社会学和人口学的研究，在其他学科的研究中非常少见，仅有1篇传播学的研究使用了该模型，这也是因为社会学和人口学更加关注社会变迁的研究；此外，已有研究的议题涉及较广，既包含社会分层、市场转型、教育问题、死亡率研究和收入研究等传统社会学和人口学议题，也有很多研究集中地讨论了当代中国民众主观态度方面的变迁。

在数据使用方面，除了1篇文献应用的是汇总数据外，其他研究利用的均是抽样调查数据。在抽样调查数据的来源方面，既包含国内机构开展的调查，也包含国外机构开展的调查；既包含如中国综合社会调查（CGSS）和中国社会状况综合调查（CSS）等多期截面数据，也包含如中国老年健康与追踪调查（CHARLS）以及中国

①文献检索的范围主要为南京大学中国社会科学评价中心发布的中国社会学科引文目录（2020-2021）中的相关期刊，时间截止到2022年1月31日。

②由于学科间的边界不够清晰，对于所在学科的判断主要通过刊载的期刊、中图文分类以及研究主题进行判定。

健康与营养调查（CHNS）等追踪数据。具体而言，在这些研究之中，中国综合社会调查数据（CGSS）的使用频次最高，为5次，同时也有研究将多源数据进行合并使用。

在模型求解方法方面，我们可以看到，HAPC模型是应用最多的方法，在20篇文献中只有2篇文献没有使用该模型。在模型选择的理由方面，梁玉成（2007）的研究指出，其选用代理变量法主要是参照前人的相关研究，在具体操作上将个体进入劳动力市场时的现代化和市场化指标作为其世代变量，将劳动力市场中特定年份的现代化和市场化指标作为其时期变量。陈心广和王培刚（2014）选择IE方法主要是他们认为IE的数学原理较为可靠，但未和其他模型（如HAPC）进行比较。高海燕等（2022）在研究过程中虽然使用了IE和HAPC两种方法，但主要使用的是IE方法而将HAPC作为稳健性检验，此外他们在叙述方法时存在不准确之处，如HAPC是在IE之后提出的，而原文颠倒了时间顺序，并且只叙述了IE方法能够得到无偏估计，忽视了大量对于IE方法批评的文献。其余的研究均选择了HAPC模型，并都在文中较为详细地阐述了模型的数学原理（实质上为多层模型的数学原理），同样也忽视了对于HAPC批评的文献。

由于APC模型不同求解方法之间存在较大的争议，吴晓刚和李晓光（2021）则提出可以借鉴国外研究的做法，采用至少两种求解方法来检验稳健性，在他们的研究中较为完整地向读者呈现出两种研究的结果。随后在高海燕等（2022）的研究中也参照了上述做法，但是仅详细报告了一种方法（IE）的结果。而我们认为，面对APC模型这种争议较大的方法，应当将两种或多种方法的结果都进行较为详细的报告。

表5　APC模型在国内社会科学研究中应用的概况

序号	文献来源	所在学科	研究主题	数据	估计方法
1	梁玉成，2007	社会学	解决传统市场转型研究模型不能分离世代效应和时期效应的问题	中国综合社会调查（CGSS）	代理变量法
2	曹惟纯、叶光辉，2014	社会学	台湾民众孝道观念的变迁	台湾社会变迁基本调查	HAPC
3	陈心广、王培刚，2014	人口学	中国城市人口年龄别死亡率的变迁	《中国卫生统计年鉴》数据	IE
4	石超、乔晓春，2017	人口学	中国人生活满意度的变迁	世界价值观调查（WVS）	HAPC
5	田丰，2017	社会学	中国农民工社会经济地位的变化	中国社会状况综合调查（CSS）	HAPC
6	范文婷等，2018	社会学	高等学历对女性教育婚配的年龄、时期和队列作用	中国妇女社会地位调查	HAPC
7	李婷，2018	人口学	中国人主观幸福感的变迁	中国综合社会调查（CGSS）	HAPC
8	王志理等，2019	人口学	中国流动人口避孕模式的变迁	全国流动人口动态监测调查	HAPC
9	吴炳义等，2019	人口学	我国老年人群高血压患病的年龄、时期和队列效应	中国老年健康影响因素跟踪调查（CLHLS）	HAPC
10	张子杨、杨慧，2019	人口学	老年妇女收入的变迁	中国妇女社会地位调查	HAPC

续表

序号	文献来源	所在学科	研究主题	数据	估计方法
11	崔凯，2019	传播学	中国民众互联网使用频率的变迁	中国综合社会调查（CGSS）	HAPC
12	石智雷等，2020	社会学	中国居民的健康（超重率和不动率）的变迁	中国健康与营养调查（CHNS）	HAPC
13	李婷等，2020	人口学	中国三代直系家庭的变迁	中国健康与营养调查（CHNS）	HAPC
14	毕文芬，2020	社会学	个人教育在就业性别观念变迁中的作用	世界价值观调查（WVS）	HAPC
15	孔国书等，2021	人口学	中国人自评健康的变迁	中国综合社会调查（CGSS）	HAPC
16	吴敏、熊鹰，2021	人口学	中国老年人口消费的变迁	中国居民收入调查（CHIP）和中国老年健康与追踪调查（CHARLS）	HAPC
17	吴晓刚、李晓光，2021	社会学	劳动力市场中过度教育的变迁	中国综合社会调查（CGSS）	C-APC & HAPC
18	王俊秀、刘洋洋，2021	社会学	中国居民隐私安全感的变化及其影响因素分析	中国社会状况综合调查（CSS）	HAPC
19	李骏，2021	社会学	城市居民的阶层认同变迁	自调查数据	HAPC
20	高海燕等，2022	社会学	中国民众社会价值观的变迁	世界价值观调查（WVS）	IE & HAPC

四、不等间距数据在APC模型中的应用

不等间距（unequal width）数据是指在年龄、时期或世代上间隔不等，例如时期间隔为5年，而年龄间隔为3年。而传统APC模型对于数据结构的要求相对严格，要符合列克西斯表（年龄时期交叉表）的分布。例如，表6中时期分别为2010年、2015年和2020年，年龄的间隔也为5年，则在时期年龄交叉表中对角线就是出生世代年份。

表6　等间距数据的APC数据结构

年龄　时期	2010	2015	2020
15~19	1989—1995	1996—2000	2005
20~24	1984—1990	1989—1995	1996—2000
25~29	1981—1985	1984—1990	1989—1995
30~34	1986—1980	1981—1985	1984—1990

但是，如果出现时期不等间距情况时，如表7中时期分别为2010年、2016年和2022年，年龄间隔依旧为5年，如果按照对角线上年份就是出生世代年份的原理来看，表6中的六条对角线分别为：1976—1980、1981—1986、1984—1992、1989—1997、1997—2002和2003—2007。很明显出现了世代重叠和缺失的情况。

表7　时期不等间距数据的APC数据结构

年龄　时期	2010	2016	2022
15~19	1989—1995	1997—2001	2003—2007
20~24	1984—1990	1992—1996	1998—2002
25~29	1981—1985	1987—1991	1993—1997
30~34	1976—1980	1982—1986	1988—1992

之所以在这里专门强调这一问题主要是因为在国内社会科学研究中时期不等间距的调查数据较为常见，一方面是因为有部分调查在执行过程中就存在不等间距情况，例如学者们使用最多的CGSS数据，目前共计公开了2003年、2005年、2006年、2008年、2010年、2011年、2013年、2015年、2017年和2018年的数据，是典型的时期不等间距的数据；另一方面，研究者为了考察更早的时期效应可能会合并其他来源的数据，例如吴愈晓等（2022）为了考察1990—2015年间性别角色观念的变迁，综合使用中国妇女社会地位调查数据和CGSS数据。然而中国妇女地位调查数据每十年进行一次，而CGSS则在2003年以后密集进行，这就势必会带来时期不等间距的情况。

对于不等间距数据，不仅会有模型本身的"不可识别"难题，而且会出现世代效应重叠导致的识别问题，例如在表6中，1981—1986、1984—1992是两个重叠的世代，他们对应的效应该如何解释成了难题。对此，霍尔福德提出，可以将年龄和时期按照二者的最小公倍数进行重新组合分组，新的年龄组和时期组由若干个原有年龄组和时期组的分组元素构成，被称为宏组，宏组内原有年龄组和时期组的元素组合被称为微组，在分析时，要将宏组和微组都纳入模型之中（详见：Holford，2006）。但不等间距数据的问题并没有引起西方学者的过多讨论，主要是因为他们的数据资源更加丰富。

但需要注意的是，如果采用传统APC模型的设定以及利用CGLM方法、代理变量法以及IE的方法则需要慎重考虑数据不等间距这一问题。对于HAPC模型而言，数据不等间距的影响依旧会存在，但对上述方法的影响相比可能相对较小，因为该模型中的世代变量已经被置于第二层，相当于一个相对独立的变量，并且杨洋和

兰德在提出HAPC模型时的示例以及很多应用都是基于不等间距数据（Yang & Land，2006，2008，2013）。相比之下，APC-I模型对于数据结构十分敏感，因为它背后的原理是完全将对角线视为世代，所以在使用APC-I模型时必须保证间隔相同，并且霍尔福德的宏组、微组分组方法可能在此模型中并不适用。这里我们可以采用一种相对粗糙的做法，即对时期进行调整，以CGSS数据为例该，数据共包含2003年、2005年、2006年、2008年、2010年、2011年、2013年、2015年、2017年和2018年共10个时期，时期间的间距几乎完全不等，我们可以将具体的时间点调整为一个时期区间，如表8所示。当然，这种做法是相对粗糙的，因为只用了某一年或两年的情况代表三年间的情况，但我们认为APC模型的关键在于刻画时间变迁的趋势，这样的处理不会产生较大的偏误。并且这可能也是唯一行之有效的办法。

表8　时期不等间距数据的调整策略

时期区间	2003—2005	2006—2008	2009—2011	2012—2014	2015—2017	2018—2019
实际包含时期	2003&2005	2006&2008	2010&2011	2013	2015&2017	2018

五、结语：APC模型的应用建议与展望

（一）APC模型的应用建议

因为年龄、时期和世代之间存在完全共线性关系，求解APC模型一直被认为是一个难以完成的任务，早在20世纪70年代就有学

者指出APC模型从本质上是无解的（Glenn，1976）。但是由于社会变迁研究以及世代效应的重要意义，许多学者不甘放弃，从理论和方法上进行了较多的研究，推动了APC模型不断发展。本文较为详细地回顾了APC模型的基本原理及其发展历程，并总结了5种最常用的求解APC模型的方法。尽管每一个方法背后都萦绕着质疑的声音，以至于出现对APC模型批评的文献比应用APC模型的文献还要多的现象，但可能正如鲍克斯和德雷伯（Box & Draper 1987）所言："实质上，所有的统计模型都是错的，只是有一些有用而已。"对待APC模型，我们不能一概地否定，也不能盲目相信。

此外，我们也可以发现，对于APC模型的求解难题与因果推断相关方法遇到的困难如出一辙。在因果推断方面，无数研究设计试图从统计方法方面来识别因果关系（Pearl，2000），同样也是遭到其他研究者的诟病。现在，学者们已经认识到，仅通过统计方法得不到因果，求解因果关系要么需要很强的理论假定（如假定不存在忽略变量），要么需要借助外力（如工具变量）或特殊的研究设计（如随机化实验和断点回归设计等），而这些外力或特殊的研究设计是否有效在本质上也需要理论假定。在APC模型方面，很多先驱者早就指出，该模型在本质上是无解的，如果一定要求解，就必须借助外力，而这个外力只能是理论（Mason et al.，1973；Mason & Smith，1985）。因而，我们认为未来的APC求解问题也可能会像因果推断系列方法的发展一样，更多地去寻求理论假定。

那么，对于现阶段研究者该如何使用APC模型？本文基于对上述5种求解方法的回顾并结合国内社会科学研究的实际情况，对于未来APC模型的应用提出如下三点建议。

第一，研究者需要明确每一种方法背后的假定，根据研究实际

进行选择。任何一个统计模型背后都会有一系列的假定，只是一些基础模型（如：回归分析等）的假定相对容易满足。而APC模型背后的假定则相对较强，但如果研究者的研究主题和数据能够满足某一种求解方法的假定，那么这种方法将会是最合适的。例如：如果研究者有充分的理由说明某两个时期（或年龄或世代）的效应会完全一致，那么CGLM方法将是最适合的；如果研究者能充分说明其他变量能够替代年龄变量（或时期变量或世代变量），那么代理变量法将会是最合适的……总之，对于APC模型的求解方法并不是越新的方法越好，而是要选择合适的。但是，去匹配每一种方法背后的假定是一个十分艰巨的任务，尤其对于IE、HAPC和APC-I等方法来说。

第二，慎重使用IE方法。建议研究者慎重使用IE方法主要有两个方面的原因：（1）IE方法背后的假定几乎难以满足，并且对于使用者而言也很难验证；（2）虽然每种方法都曾被批评，但是IE方法自提出之后就受到西方学术界更为严苛的批评，例如在权威期刊《人口学》（*Demography*）2013年第6期上就刊登了五篇由APC领域的顶尖专家撰写的文章（Held & Riebler，2013；Luo，2013a，2013b；O'Brien，2013；Yang & Land，2013），其中除了杨洋和兰德（IE方法的提出者）之外，其他四篇文章都一致认为IE方法存在重大缺陷，即在APC领域的顶尖学者眼中，IE是有很大问题的。

第三，在应用APC模型时应进行多种策略的稳健性检验。吴晓刚和李晓光（2021）曾建议在使用APC模型时应采用两种以上的求解方法进行稳健性检验。但本文认为，由于APC模型的争议较大，对于APC模型的稳健性检验应不止于使用不同的求解方法进行检验，因为也是有几率出现两种求解方法均是错误的情况。因而，要

借鉴经济学的传统，采用更多的策略进行检验，如替换研究的因变量、替换研究的数据或单独截取部分年份数据进行检验。

（二）APC模型在国内应用的机遇与挑战

虽然现阶段有关APC模型的讨论充斥着各种各样的质疑，但我们依然坚信在未来，APC模型在国内社会科学的研究还是充满机遇的，当然也会面临一些挑战。首先，社会变迁是一个永恒的研究主题，也是具有高度理论和现实意义的议题。尤其是中国社会，我们自20世纪初以来经历了急剧的社会变迁，从战争、饥荒到新民主主义革命，新中国成立以后，也经历了计划经济时期、改革开放和走进新时代等诸多重要的历史阶段。急剧的社会变迁如何影响个体的生活机遇？或，在历史的长河中个体的特征、行为方式、态度和观念是如何改变的？所有这些重要问题的回答可能都需要借鉴世代分析的理念和APC的方法和模型。

其次，在数据方面，我国的社会调查和数据共享工作虽然起步较晚，但是经过多年的积累已经具备一定的规模，为APC模型的应用提供了机遇。例如：社会科学研究经常使用的中国综合社会调查（CGSS）从2003年开始已经有近20年的时间，已经公布了11期数据。中国社会状况综合调查（CSS）从2006年开始，已经公布了7期数据，以及中国健康与营养调查（CHNS）、世界价值观调查（WVS）等数据，都为APC模型分析提供了充足的经验资料。但是，在数据方面也存在两个方面的挑战：其一是调查前后的测量不一致问题突出，虽然我们有了较好的数据积累，但是对很多概念的测量面临在不同年份不一致的尴尬情况，或没有进行持续的调查，这无疑会限制了APC模型的应用；其二是我国很多调查数据存在调查年

份的不等间距情况，例如CGSS调查共有2003、2005、2006……在部分年份并不是规律的间隔一年或两年，这也为APC模型的使用带了一定的困难，但针对这一问题目前已有学者提出应对策略，前文已经介绍。

最后，有关APC模型的“不可识别”难题的应对，目前学界还没有共识，这的确会为应用研究带来挑战，选择何种方法求解APC模型是摆在每一位研究者面前的难题。但是，我们不能因为对其中某一种方法的质疑，而否定利用该种方法的研究结果，因为每一种方法都具有一定的合理性。而正是因为APC模型还处于发展阶段，也为社会科学研究者提供了一个重要的机遇。目前有关APC方法的讨论与开拓主要集中在西方学界，更多是为了理解西方社会的变迁模式。自20世纪初期以来，中国经历了急剧的社会变迁，当前还处于社会的转型时期。由于社会、经济和文化等诸多方面的独特性，中国社会变迁的模式，尤其是时期、世代和年龄三个时间维度的变化趋势和机制可能有别于西方社会。因此，除了借鉴已有的APC模型应用于我国的数据资料之外，也希望国内的学者能结合中国社会变迁的实践，在APC模型求解方法的开拓上贡献更多的“中国智慧”!

参考文献

埃尔德，2002，《大萧条的孩子们》，江苏，南京：译林出版社。

毕文芬，2020，《个人教育在就业性别观念变迁中的作用：基于HAPC模型的分析》，《深圳大学学报（人文社会科学版）》第5期。

曹惟纯、叶光辉，2014，《高龄化下的代间关系——台湾民众孝道信念变迁趋势分析（1994-2011）》，《社会学研究》第2期。

沈其君，1992，《生命统计资料中年龄和时期趋势效应的估计》，《中国卫生统计》第1期。

陈心广、王培刚，2014，《中国社会变迁与国民健康动态变化》，《中国人口科学》第2期。

崔凯，2019，《后喻文化与双重滞后：中国互联网的十年扩散历程（2005～2015）——基于4组CGSS数据的APC模型分析》，《现代传播（中国传媒大学学报）》第2期。

范文婷、宋健、李婷，2018，《高等教育与女性婚配：基于年龄、时期和队列的视角》，《人口学刊》第2期。

高海燕、王鹏、谭康荣，2022，《中国民众社会价值观的变迁及其影响因素——基于年龄-时期-世代效应的分析》，《社会学研究》第1期。

卡尔·曼海姆，2002，《卡尔·曼海姆精粹》. 南京：南京大学出版社。

孔国书、惠长虹、李路路，2021，《中国居民自评一般健康的队列差异研究——兼论“人口红利”的健康效应》，《人口学刊》第6期。

李婷、宋健、成天异，2020，《中国三代直系家庭变迁：年龄、时期、队列视角的观察》，《人口学刊》第3期。

李骏，2021，《从收入到资产：中国城市居民的阶层认同及其变迁——以1991-2013年的上海为例》，《社会学研究》第3期。

李婷，2018，《哪一代人更幸福?——年龄、时期和队列分析视角下中国居民主观幸福感的变迁》，《人口与经济》第1期。

梁玉成，2007，《现代化转型与市场转型混合效应的分解——市场转型研究的年龄、时期和世代效应模型》，《社会学研究》第4期。

石智雷、顾嘉欣、傅强，2020，《社会变迁与健康不平等——对第五次疾病转型的年龄—时期—队列分析》，《社会学研究》第6期。

石超、乔晓春，2017，《中国人生活满意度的年龄-时期-队列效应分析》，《人口与发展》第4期。

苏晶晶、彭非，2014，《年龄-时期-队列模型参数估计方法最新研究进展》，《统计与决策》第23期。

田丰，2017，《逆成长：农民工社会经济地位的十年变化（2006-2015）》，《社会学研究》第3期。

王俊秀、刘洋洋，2021，《中国居民隐私安全感的变化及其影响因素——基于年龄—时期—队列分析》，《江苏行政学院学报》第6期。

王志理、张婧文、庄亚儿，2019，《2011-2017年中国流动人口避孕模式变化分析》，《人口学刊》第5期。

吴愈晓、王金水、王旭洋，2022，《中国性别角色观念变迁（1990-2018）：年龄、时期和世代效应及性别差异模式》，《中华女子学院学报》第4期。

吴炳义、董惠玲、于奇等，2019，《生命历程视角下老年人口高血压患病的年龄-时期-队列效应分析》，《人口与发展》第3期。

吴敏、熊鹰，2021，《年龄、时期和队列视角下中国老年消费变迁》，《人口与经济》第5期。

吴晓刚、李晓光，2021，《中国城市劳动力市场中教育匹配的变迁趋势——基于年龄、时期和世代效应的动态分析》，《中国社会科学》第2期。

张子杨、杨慧，2019，《老年妇女收入双重劣势研究——以1926-1950年出生队列为例》，《人口与发展》第5期。

Bell, A. & K. Jones 2015, "Should age-period-cohort analysts accept innovation without scrutiny? A response to Reither, Masters, Yang, Powers, Zheng and Land." *Social Science & Medicine* 128, 331-333.

Box, G. E. P. & N. R. Draper 1987, *Empirical model-building and response surfaces*. Oxford, England: John Wiley & Sons: xiv 669.

Fosse, E. & C. Winship 2019, "Analyzing Age-Period-Cohort Data: A Review and Critique." *Annual Review of Sociology* 45(1), 467-492.

Frost, W. H. 1939, "THE AGE SELECTION OF MORTALITY FROM TUBERCULOSIS IN SUCCESSIVE DECADES." *American Journal of Epidemiology* 30-SectionA(3), 91-96.

Giele, J. Z. & G. H. Elder 1998, *Methods of life course research: qualitative and quantitative approaches*. Thousand Oaks, Calif: Sage Publications.

Glenn, N. D. 1976, "Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period and Cohort Effects." *American Sociological Review* 41(5), 900-904.

Glenn, N. D. 2005, *Cohort analysis*. Sage.

Heckman, J. & R. Robb 1985, "Using Longitudinal Data to Estimate Age, Period and Cohort Effects in Earnings Equations." Mason W. M., Fienberg S E, eds.//, *Cohort Analysis in Social Research: Beyond the Identification Problem*. New York, NY:

Springer: 137-150.

Holford, T. R. 2006, "Approaches to fitting age-period-cohort models with unequal intervals." *Statistics in Medicine* 25(6), 977-993.

Kahn, J. R. & W. M. Mason 1987, "Political Alienation, Cohort Size, and the Easterlin Hypothesis." *American Sociological Review* 52(2), 155-169.

Kupper, L. L., J. M. Janis & A. Karmous et al. 1985, "Statistical age-period cohort analysis: A review and critique." *Journal of Chronic Diseases* 38(10), 811-830.

Luo, L. & J. S. Hodges 2020a, "Constraints in Random Effects Age-Period-Cohort Models." *Sociological Methodology* 50(1), 276-317.

Luo, L. & J. S. Hodges 2020b, "The Age-Period-Cohort-Interaction Model for Describing and Investigating Inter-cohort Deviations and Intra-cohort Life-course Dynamics." *Sociological Methods & Research* 51(3).

Mason, K.O., W. M. Mason & H. H. Winsborough et al. 1973, "Some Methodological Issues in Cohort Analysis of Archival Data." *American Sociological Review* 38(2), 242-258.

Mason, W. M. & H. L. Smith 1985, "Age-Period-Cohort Analysis and the Study of Deaths from Pulmonary Tuberculosis." Mason W. M., Fienberg S E, eds.//, *Cohort Analysis in Social Research: Beyond the Identification Problem*. New York, NY: Springer: 151-227.

O'Brien, R. M. 2011, "Constrained Estimators and Age-Period-Cohort Models." *Sociological Methods & Research* 40(3), 419-452.

Pearl, J. 2000, *Models, Reasoning, and Inference*. Cambridge, UK: Cambridge University Press.

Rodgers, W. L. 1982, "Estimable Functions of Age, Period, and Cohort Effects." *American Sociological Review* 47(6), 774-787.

Ryder, N. B. 1965, "The cohort as a concept in the study of social change." *American sociological Review* 30(6).

Winship, C. & D. J. Harding 2008, "A Mechanism-Based Approach to the Identification of Age-Period-Cohort Models." *Sociological Methods & Research* 36(3), 362-401.

Yang, Y. 2008, "Social Inequalities in Happiness in the United States, 1972 to 2004: An Age-Period-Cohort Analysis." *American Sociological Review* 73(2), 204-226.

Yang, Y., W. J. Fu & K. C. Land 2004, "A Methodological Comparison of Age Period-Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models." *Sociological Methodology* 34(1), 75-110.

Yang, Y. & K. C. Land 2006, "A Mixed Models Approach to the Age-Period Cohort Analysis of Repeated Cross-Section Surveys, with an Application to Data on Trends in Verbal Test Scores." *Sociological Methodology* 36(1), 75-97.

—2008, "Age-Period-Cohort Analysis of Repeated Cross Section Surveys: Fixed or Random Effects?" *Sociological Methods & Research* 36 (3), 297-326.

—2013, *Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications.* Taylor & Francis.

Bell, A. & K. Jones 2014, "Another 'futile quest'? A simulation study of Yang and Land's Hierarchical Age-Period-Cohort model." *Demographic Research* 30, 333-360.

Held, L. & A. Riebler 2013, "Comment on 'Assessing Validity and Application Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort(APC)Problem'." *Demography* 50(6), 1977-1979.

Luo, L. 2013a, "Paradigm Shift in Age-Period-Cohort Analysis: A Response to Yang and Land, O'Brien, Held and Riebler, and Fienberg." *Demography* 50(6), 1985-1988.

Luo, L. 2013b, "Assessing Validity and Application Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort Problem." *Demography* 50(6), 1945- 1967.

O'Brien, R. M. 2013, "Comment of Liying Luo's Article, 'Assessing Validity and Application Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort Problem'." *Demography* 50(6), 1973-1975.

Yang, Y. C.& K. C. Land 2013, "Misunderstandings, Mischaracterizations, and the Problematic Choice of a Specific Instance in Which the IE Should Never Be Applied." *Demography* 50(6), 1969-1971.

Charles, F. Manski 1995, *Identification Problems in the Social Sciences.* Cambridge, Mass: Harvard University Press.

孙维茁，张立．定性比较分析（QCA）综述：内涵、操作与应用[M/OL]//赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

定性比较分析（QCA）综述：内涵、操作与应用

孙维茁　张立[①]

摘要：定性比较分析（QCA）是一类研究思想和分析技术的总称。作为研究思想，其代表了集合论和布尔代数体现的组态视角，意味着将每个案例都视作整体，通过案例间的对比分析，得到不同条件和条件组合对结果的解释程度；作为分析技术，其代表了cs-QCA、fs-QCA、mv-QCA、T-QCA等多样的方法工具。研究实践证明，QCA的优势与不足并存，机遇与挑战同在。本文对其理论和方法内涵、操作步骤与策略、当前的应用情况进行了梳理，有利于在研究中培养新思维和做出新尝试。但QCA仍伴随诸多不足与质疑，其方法完善和应用前景等仍有待在未来研究中进一步加强。

关键词：定性比较分析；社会研究方法；组态分析；QCA

①作者简介：孙维茁，西安交通大学新闻与新媒体学院硕士研究生，主要研究兴趣为新媒体与社会治理、媒介化、舆情治理，联系方式：sunweizhuo@foxmail.com；张立，西安交通大学新闻与新媒体学院副教授、硕士研究生导师，主要研究兴趣为新媒体与基层社会治理、风险传播、舆情生态治理，联系方式：zhanglixinwen@mail.xjtu.edu.cn。

Abstract: Qualitative Comparative Analysis (QCA) is a general term for a class of research ideas and analytical techniques. As a research idea, it represents set theory and Boolean algebra reflected in the configuration perspective, which means that each case is regarded as a whole. Through comparative analysis between cases, the explanation degree of different conditions and combinations can be obtained. As an analytical technique, it represents various operation methods such as cs-QCA, fs-QCA, mv-QCA and T-QCA. Research practice has proved that QCA has both advantages and disadvantages, as well as opportunities and challenges. In this paper, the connotation of its theory and method, operation steps and strategies, and the current application are sorted out, which is conducive to cultivating new thinking and making new attempts. However, QCA is still accompanied by many shortcomings and doubts, and the perfection of the method and the exploration of its application need to be further strengthened in the future research.

Key words: Qualitative Comparative Analysis; Social Research Method; Configuration Analysis; QCA

一、引言

定性比较分析（Qualitative Comparative Analysis，简称QCA）是用于社会科学研究的一种中小样本案例组态比较分析的思想和方法，组合了定性和定量策略，同时又兼顾了量化和质化的双重取向，被学界喻为方法上的“第三条道路”。

1987年，美国社会学者查尔斯·拉金（Charles C.Ragin）的著作《比较方法：走向超越定性与定量策略之路》（*The Comparative*

Method: Moving Beyond Qualitative and Quantitative Strategies）出版，作者试图将定性研究和定量研究的最佳特征结合起来，开拓两种研究的中间道路——QCA。随后，拉金又于2000年出版了《模糊集社会科学》（*Fuzzy-set Social Science*），2008年出版了《重新设计社会探究：模糊集及其超越》（*Redesigning Social Inquiry: Fuzzy Sets and Beyond*）。据此，定性比较分析方法逐渐发展成熟，并在政治学、经济学、社会学等多个社会科学领域得到应用。

关于QCA的研究在西方学界成果比较丰富。（1）作为研究工具的QCA，学者们主要对其具体类型进行介绍，包括清晰集（cs-QCA）（Grofman & Schneider，2009）、模糊集（fs-QCA）（Longest & Vaisey，2008）、多值集（mv-QCA）（Vink & Van Vliet，2009）和时序集（T-QCA/TS-QCA）（Caren & Panofsky，2005）。此外，学者们也关注到了QCA在实际操作中的案例选择标准（Skaaning，2011）和结果稳健性检验（Schneider & Wagemann，2010）等问题。（2）在QCA的应用方面，学者们一方面提到了对服务创新配置（Ordanini，Parasuraman & Rubera，2013）、民主社会治理（Schneider & Wagemann，2006）等各种社会现象进行分析时的应用，另一方面提到了对公共政策（Rihoux，Rezsöhazy & Bol，2011）、组织管理（Seny Kan et al.，2016）等特定领域的相关应用。

相比较而言，国内学者的研究尚处于探索阶段。2009年，QCA方法被用于分析互联网对业主集体抗争的影响，通过4个解释变量和15个案例，验证了互联网的动员效用（黄荣贵、桂勇，2009），因作者在研究中未对QCA方法进行介绍，所以在理解上具有难度。2017年，国内学者对QCA相关国外著作进行内容整合，并翻译出版《QCA设计原理与应用：超越定性与定量研究的新方法》。管理

学（杜运周、贾良定，2017）、新闻传播学（毛湛文，2016）、国际关系学（刘丰，2015）等学科从宏观角度进行特定领域的方法引入；此外，回顾展望类的研究也逐年增多。但是，总体来看，研究的全面性、系统性略显不足。

本文将对QCA的内涵、操作与应用进行梳理和介绍，以期为相关研究提供参考。

二、QCA的历史演进

要实现对QCA内涵、操作、应用等方面的系统了解和整体把握，将其置于历史中进行回顾十分必要。QCA的逻辑基础在于通过比较，以厘清复杂现象背后的因果关系。[①] 20世纪80年代末90年代初，QCA开始被广泛应用于社会科学研究，自该方法诞生起，研究者们就致力于用其实现定性与定量研究的优势整合，即融合案例导向与变量导向两种思路（Ragin，1987：84）。在逻辑方面，作为对“经常连结”质疑的回应，QCA发展出新的因果关系概念，即“多重并发因果关系”（multiple conjunctural causality）：（1）“多重”是指引发某一结果的路径组合丰富；（2）“并发”是指构成每条路径的条件组合各异；（3）“等效性”是指不同因果路径会引发相同结果。在实践方面，QCA充分结合了案例内分析（within-case anal-

①相关思想最早可以追溯至18世纪，休谟（David Hume）对因果关系的真实性和必然性进行反驳。休谟提出“经常连结”（constant conjunction）概念并指出其中不合理之处（Hume，1999：74-81），主张打破惯常的因果思维；密尔（John Stuart Mill）的求同法和求异法则为这一思想提供了可操作路径，即通过匹配和比较被调查对象，发现使其在某方面维持一致或产生差异的单一原因（Mill，2002：390），然而这种设计过于理想，仅依靠单纯的观察与比较，难以囊括所有要素或排除其他要素，但至少为QCA的完善提供了可靠思路。

ysis）和跨案例分析（cross-case analysis）的操作过程。[①]经过多年演进，QCA不断拓展因果关系的分析框架，在认识论和方法论上实现突破：其一，可加性假设被打破，单个原因不再对结果产生独立影响；其二，不具有唯一确定的因果关系；其三，因果效应间不具有正向或负向影响的绝对一致性；其四，因果关系不具有对称性，多因可能解释一果（里豪克斯、拉金，2017：7-8）。

从最初的设计构想来看，QCA至少能应用于五种情形：（1）直接以真值表的形式呈现，为研究者提供汇总数据、描述案例的工具；（2）借前因相同但结果不同的矛盾组态检验数据一致性，通过与案例不断“对话”进行优化分析；（3）以分析结果证实或证伪理论与假设，并明确改善方向；（4）快速检验没有先验理论或模型的初步猜想；（5）通过不含矛盾组态的真值表和简化解获得新认识，为发展新理论奠定基础（里豪克斯、拉金，2017：14-15）。

三、作为研究思想和分析方法的QCA

查尔斯·拉金在提出定性比较分析时，赋予了其技术和方法双重角色（Rihoux & Lobe，2009），后来的研究大多沿用这两种分析角色。有学者将其解读为研究取向和操作技术（毛湛文，2016），也有学者认为是技术集和方法（Rihoux et al.，2011），虽然解读的名称不同，但内涵基本一致。技术指向通过集合论和布尔代数体现的组态分析思想和视角，方法指向在具体操作过程中

①由于因果关系依赖组态与情景，必须首先将各案例作为独立整体进行全面分析，并在这一环节中用变量赋予案例客观性与科学性，以兼顾定义关键变量的定量特征和保持整体性视角的定性特征，再在所有案例间进行计算与分析，值得注意的是，用于计算的案例均同等重要，分析得到的组合路径和因果关系因此更为多样。

选取的分析软件和研究工具，本文将之概括为研究思想和分析方法，进行分别阐述。

（一）作为研究思想的QCA

组态分析和集合论思想是QCA的核心逻辑。组态分析最早在组织战略研究领域提出，西方学者认为组织形态（organizational configuration）是共同发生的、由概念上可区分的特征构成的多维特征群（Meyer，Tsui & Hinings，1993），并在此基础上发展出了组态分析视角，强调相互关联的结构和整体集群，而非分散或单元个体的结合实体（Fiss，2007）。因此，组态分析（organizational analysis）思想就是研究整体和系统层面的案例组态作用，而非研究单个变量对结果的作用，这一思想在提出后的很长一段时间内并未得到充分重视和广泛应用，主要原因在于空有理论，缺少方法论的配合，拉金提出的QCA填补了这一空白。拉金借用数学中的集合思想，把社会科学领域的大部分命题都用集合的隶属关系来表示，将集合论思想和布尔代数逻辑引入了社会科学研究领域。

如果将某个研究问题或社会现象作为集合，构成这一问题或现象的原因就是不同的子集，在这样的集合思想上，QCA利用布尔代数的运算法则对比多案例，并在不同集合间分析因果关联，用数学思想解决逻辑问题。[①]此外，为了解决复杂社会现象中多案例的矛

①广义的布尔代数是一套基于二值的表示逻辑推理的符号系统，其用字母表示变量，最基本的逻辑值为0和1，0代表假（FALSE），即变量不出现，1代表真（TRUE），即变量出现，不同变量间存在“非（NOT）”的一元逻辑算法，“和（AND）”“或（OR）”的二元逻辑算法三种基本运算（程代展、赵寅、徐相如，2011）。

盾组合，QCA运用还要遵循“布尔最简化”原则。[①]

表1 社会科学研究传统统计方法与QCA思想对比(Ragin,2008)

传统统计方法	QCA
变量	集合
测量	校准
因变量	质化结果
特定样本	构造样本
相关性	设定理论关系
相关矩阵	真值表
净效应	因果食谱(Recipe)
反事实估计	反事实分析
概率推断	比较判断
单一因果模型	多样因果模型

（二）作为分析方法的QCA

随着组合分析思想和研究工具的逐渐成熟，运用QCA分析的研究问题和社会现象日趋复杂，早期的二分变量已经不能满足研究需求，拉金和其他学者发展出了清晰集（cs-QCA）、模糊集（fs-QCA）、多值集（mv-QCA）和时序集（T-QCA/TS-QCA）等多样的研究工具。

1. 清晰集(crisp set):cs-QCA

清晰集是指仅用0和1对变量校准（calibrate）并进行QCA分析

①即当两个布尔代数表达式中仅有一个条件不同，而它们得出了同样的结果，则该条件可以删除，得到精简后的布尔代数表达式（Ragin，1987：26）。

的方法，是初代分析工具。布尔代数最基本的逻辑值是0和1，因此在QCA的变量校准中，无论是解释变量还是结果变量，0/1都是对变量存在与否的基础划分，相当于logistic回归（Grofman & Schneider，2009）。由于“是与否”的变量判断相对清晰，解释变量和结果变量存在明显的对应关系，明确的判断标准尤为重要，而整个过程的操作也相对简单。

清晰集工具多适用于相对简单、关系明确和可以二分的问题或现象，例如流行语高热度或低热度传播（周俊、王敏，2016）、个人产生行为与否（Mozas-Moral，2016）、政府创新扩散自上而下或自下而上（刘晓亮、侯凯悦、张洺硕，2019）、企业跨国并购成功或失败（程聪、贾良定，2016）等。

2. 模糊集(fuzzy set):fs-QCA

模糊集是指用［0，1］之间的值对变量校准并进行QCA分析的方法。在分析复杂现象时，很多条件难以简单归为存在或不存在，因此学者提出了“隶属度”（membership scores）的概念，即某条件在多大程度上隶属于某集合。相应地，由研究者根据每个变量与0或1代表情况的差距，决定运用模糊集的定性锚值（qualitative anchor）[①]，做进一步校准和分析。

模糊集工具多适用于情况复杂、模糊性强、难以二分的问题或现象，可以用于评估交叉、包含等集合关系，例如研究者讨论抗争诉求的实现程度时，将诉求实现却付出重大代价的“惨胜”校准为

①常见的有三值锚值（0/0.5/1）、五值锚值（0/0.25/0.5/0.75/1）、六值锚值（0/0.2/0.4/0.6/0.8/1）等（Longest & Vaisey，2008），值得注意的是，锚值并不要求均分，研究者可以根据需要进行划分，比如0/0.8/1。

0.6（黄荣贵、郑雯、桂勇，2015）、采用三值锚值划分美国介入领土争端的干涉程度（迟永，2014）、发展两套三值锚值模糊集校准标准评判组织绩效的高低（Fiss，2011）等。

3. 多值集(multi-value):mv-QCA

多值集是指用0和1以外的值对变量校准并进行QCA分析的方法。不同于多用于衡量程度的模糊集，多值集在分析类别明确的问题或现象时具有优势，与模糊集明显的区别在于变量条件之间仍然是离散（discrete）的（Vink & Van Vliet，2009），而且各条件往往有顺序或等级类别，类似于量化研究中的定类变量，由研究者根据需要确定分类的数量。现有的多数研究都进行了三类划分，即0、1、2，超越了二元变量的范围。最早提出多值集思想的学者认为交通信号灯是应用多值集的典型情境，将红灯、黄灯、绿灯分别校准为0、1、2（Cronqvist，2004）。

多值集工具多适用于可以进行明确分类的问题或现象，例如有研究者将公民参与政策制定程度分为象征性参与、协商性参与和完全性参与（王洛忠、崔露心，2020）、将政策试点结果分为全国推广、局部推广和政策夭折（陈宇、闫倩倩，2019）等。

4. 时序集(Temporal/Time Series):T-QCA/TS-QCA

时序集是指将时间维度嵌入QCA条件变量的因果分析中，T-QCA强调条件发生作用的顺序，TS-QCA强调条件在某时间内发生的变化。前三种分析工具都只强调了对条件变量的细分或拓展，而假定时序变化不会影响结果，QCA也因此遭受了一些质疑，有学者认为其未能展现变量在因果路径中真正的变化方式，也无法展现案

例的历时性变化（释启鹏，2016），例如，A*B 和 B*A 的组合被认为是一样的。随后，有学者发展了时序集工具，T-QCA 要求研究者明确事件展开的顺序，以及对影响时序性的条件保持敏感（Caren & Panofsky，2005），TS-QCA 分析跨时间维度的变量，更加具有动态性，提出者用 QCA 分析截面数据和时间序列数据，并分为汇总型（pooled QCA）、固值型（fixed effects QCA）、时差型（time differencing QCA）（Hino，2009），用“—”表示条件作用的顺序，即 A—B 和 B—A 具有不同的意义。举例来说，如果要研究某事件爆发的原因，前三种 QCA 工具只能分析导致事件爆发的因素组合，而 T-QCA 可以分析不同因素的组合顺序对事件爆发的影响，TS-QCA 则可以分析几年间影响事件爆发的因素变化。

时序集工具多适用于具有动态性、时间属性（顺序或变化）明显并对结果产生影响的问题或现象，例如预测趋势、研究发展进程、深入分析影响因素等，但由于概念提出较晚、工具不够成熟等原因，目前尚未检索到相关应用文章。

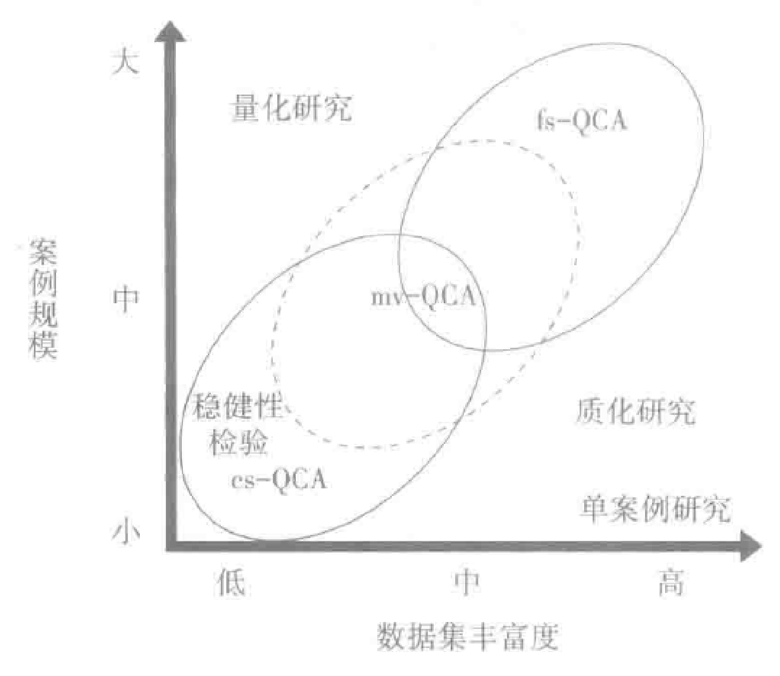


图 1　部分 QCA 工具适用图（Rihoux，2006）

四、QCA的操作步骤

（一）案例选择

定性比较分析的本质是基于案例（case-based）的研究，它要求将每个案例看作一个实体，并且在过程中不破坏其完整性，即在案例范围内分析各条件变量对结果的影响作用（Marx, Rihoux & Ragin, 2013），契合前文提及的案例内分析与跨案例分析相结合的思路。在QCA的案例选择上，主要需注意两点：数量和标准，同时要处理好案例与理论的关系。

在最早的研究中，拉金为中等规模样本研究设计了QCA，在后期发展中，社会科学研究相对宏观与中观的特征促使学者们多选取中小规模样本进行研究，学界逐渐达成了案例数量要偏少，且和条件数量平衡的共识（Marx & Dusa, 2011），但并未形成统一的数量标准。有学者认为案例少于10个为小规模样本，10—50个为中等规模样本，超过50个为大规模样本，并通过研究发现22—23个案例是大多数QCA研究使用的样本规模（Rihoux et al., 2013）；也有学者认为15—80个案例都可以用于QCA研究，但20—40个是理想状态（万筠、王佃利，2019）。与此同时，部分研究者将QCA创新性地引入了大规模样本研究，以期为实证研究提供新的视角和理论建构，但可能出现更大程度上的误差（Fiss, Sharapov & Cronqvist, 2013）。目前，越来越多的研究尝试进行大规模样本的探索和运用，但中小规模样本仍是主流，在具体数量上没有统一规定。

由于QCA的研究样本规模偏小，在案例选择上更需慎重。一

方面，国外学者对案例本身应具有的特征提出了较为全面的见解。一种观点认为案例的内在重要性、独立性、内部证据、逻辑性和代表性是选择标准（Gerring，2007）。内在重要性是指案例感知重要，有明显的历史地位，引起过学界、社会的广泛重视；独立性是指案例间彼此独立，是一个相对有边界的单元，不会相互影响而使结果产生偏差；内部证据是指各案例本身存在充足、可靠、可获得的数据支撑，使研究案例有价值；逻辑性是指研究案例逻辑自洽，在客观的案例情况和主观的研究人员判断上都是令人信服的样本；代表性是指案例典型，可覆盖的范围广，能够代表某一研究问题或社会现象。还有观点认为研究者应确保案例共享充足的背景特征、通过案例可以获得结果的明确定义、案例同时涵盖积极和消极结果，同时保证对案例内容及案例类型或种类的熟悉度（Rihoux & Ragin，2009）。

另一方面，从QCA的适用情况和结果侧重来看，其操作分析依靠案例间的比较，而多重并发因果导致的复杂社会议题是其研究的重点，因此案例间必须具有一定的可比较性，既要有共同背景或特征带来的相似性，以使案例比较能够进行，也要有在最少案例中拥有最多特质的充分多样性，以使案例比较具有意义（里豪克斯、拉金，2017：18-19），做到“最大相似”与“最大差异”相结合的系统设计（Przeworski & Teune，1970：39）。可见，学界普遍认为案例选择需要有明确且详细的标准，但在具体的标准内涵上，研究者持有不同的观点，主要应考虑案例的显著、多样和支撑材料的全面、可得，还要平衡案例复杂性和分析的可操作性。此外，在研究过程中，还应动态增减案例，而非将其作为固定的样本。

通过QCA，研究者能够在案例与理论间实现“对话”，因此必

须处理好案例与理论的关系。在研究的初期阶段，分析模型的建立和案例的选择都有赖于理论基础，同时也要以作为“例外情况”或“异常值”的案例对理论进行补充；在分析过程中，掌握有关QCA思想和逻辑能够帮助研究者将变量操作化，并解决一些看似矛盾的案例情况，推动对案例的重新审视和选取；在结果讨论时，理论有利于不同路径的重新整合与再归纳，从中提炼出最具代表性和解释力的组合方式并提供解释、获得对案例的整体认识，而相对突出的个案则可以为理论发展指引方向（里豪克斯、拉金，2017：6-7）。

（二）变量/条件确定

案例可操作化的关键和首要一步是确定解释变量和结果变量。结果变量主要由研究内容和QCA的具体分析方法决定，解释变量以已知研究成果和文章理论框架为基础，多被表述为条件，在条件确定上，主要需注意数量和策略。

将条件数量维持在中等水平是学界的共识。有学者分析近三十年的QCA相关文献发现，使用不同分析方法的QCA研究在条件数量设置上几乎没有差异，平均设置条件数为5个，而常用的数量范围为4—6个（Rihoux et al.，2013）。条件并非越多越好，原因在于数量的增多会带来作用强度（strength）和方向（direction）上的不稳定，一方面，逻辑余数（logical reminders）会大幅增长，即出现案例未体现却实际存在的组合，也就是所谓的有限多样性（limited diversity）问题；另一方面，分析过程和结果的复杂性都会增加，使程序运行和理论解释更为艰巨。但是也有少数学者认为条件数量保持中等的论断存在局限，主要在于微观分析需要比宏观分析更加异质（heterogeneous）和丰富（numerous）的数据。

同时，条件增加在一些情况下能够减少矛盾行（contradictory rows）或提高一致性值（consistency values），对构建真值表和结果分析有积极作用（Schneider & Wagemann，2010）。

在变量的确定策略上，有学者总结出了五种方法：综合法（comprehensive）、视角法（perspectives）、显著性法（significance）、二次观察法（second look）和组合法（conjunctural）（Amenta & Poulsen，1994），还有学者在此基础上附加了归纳法（inductive）（Rihoux & Ragin，2009）。综合法是指依靠所有现存的理论、假设和解释提出变量；视角法是指从实证文献的主要理论视角提供混合变量；显著性法是指采用在推论统计分析中具有显著系数或效应的变量；二次观察法是指将无关变量加入，观察其被拒绝的因素是否会对研究结果产生复杂影响；组合法是指变量的确定由具有组合性质的解释决定；归纳法是指基于案例知识，而非现有的理论确定条件。变量确定一般应注意：（1）可以反映所有案例的某一方面，且必须是变化的条件；（2）控制数量保持较低水平；（3）平衡案例数量和条件数量之间的关系；（4）建立与结果相关的假设，最好以充分或必要的形式表述（Rihoux & Ragin，2009）。

研究者需明确各方法的优势和弊端，以选择适合的策略。综合法覆盖面广，但可能会导致分析过程中优先考虑变量间的理论联系，而非实际经验，产生较多无意义组合；视角法回应了综合方法的弊端，与标准的假设检验相似，可以得到较为真实的案例和组合情况，但在评估理论方面存在缺陷，从多个视角中提取出个别变量可能会破坏理论的完整性和效用性，且不易整合成组合；显著性法有利于控制变量的数量，但并不完全符合QCA的组合认识论，即只选择显著的变量有失异质性和全面性；二次观察法可以使研究者

更加优先和客观地排除不可取的独立因素，使某因素在一些案例中缺失，而在另一些案例中存在，但这要求研究者可以预判组合结果，否则可能会过早地排除有关因素；组合法被一些学者认为更适用于QCA研究，在于它可以产生因果异构的组合结果，能够以组合形式验证假设、建构理论和分析案例。尽管具体的策略仍由研究者根据需要确定，但不论使用何种方法，变量确定都应以前期丰富的理论知识和研究过程中获得的实证发现为基础。

（三）校准与真值表构建

要用QCA软件处理分析数据，必须将数据根据变量集合中的隶属度集合校准（calibrate）为可以进行QCA分析的值。集合校准的前提是确定定性锚值[1]。数据校准后，得到解释变量和结果变量的所有组合和对应值，通过为各变量进行字母编码，并将这些组合以表格的形式呈现出来，得到真值表（truth table）。[2]

---

①即选择前文提到的一种操作方法，根据研究内容划分锚值点（anchor point），并在详细阐述理论和结合实践经验的基础上，确定各数据高于或低于该锚值点，从而赋予其隶属度，完成校准。在具体的操作步骤上，可以分为四步：（1）明确各条件和结果，即解释变量和结果变量的内涵；（2）将数据与变量对应，进行分类，并确保类别可以全面、客观地反映变量；（3）确定集合精准度，即定性锚值，设置各变量集合的锚值点并阐释意义，区分各案例数据在集合中的隶属度；（4）为数据进行赋值校准和过程中的修正（Basurto & Speer，2012）。

②在真值表的构建中，有五种组合呈现的形式："1"代表组合情况存在、"0"代表组合情况不存在、"—"代表组合情况不确定、"C"代表组合情况矛盾，即"0"和"1"同时存在、"L"或"R"代表组合存在逻辑余数，即有研究案例未体现却实际存在的组合情况。一种组合呈现形式即为一种组态，每种组态可对应零个、一个或多个案例，反映在结果变量中。研究者应通过重选案例、修正条件、更改工具或方法等尽力避免"—"和"C"的出现，由于个体研究存在局限性，"L"或"R"的情况难以避免，但也可以为研究者提供新的方向（Rihoux & Ragin，2009）。

表 2 真值表(虚拟数据)

解释变量(原因条件)				结果变量	案例数
A	B	D	E	F	
1	0	1	1	C	3
0	0	1	0	C	4
1	1	1	1	C	4
0	1	1	1	1	1
0	0	0	0	0	5
1	0	1	0	1	1

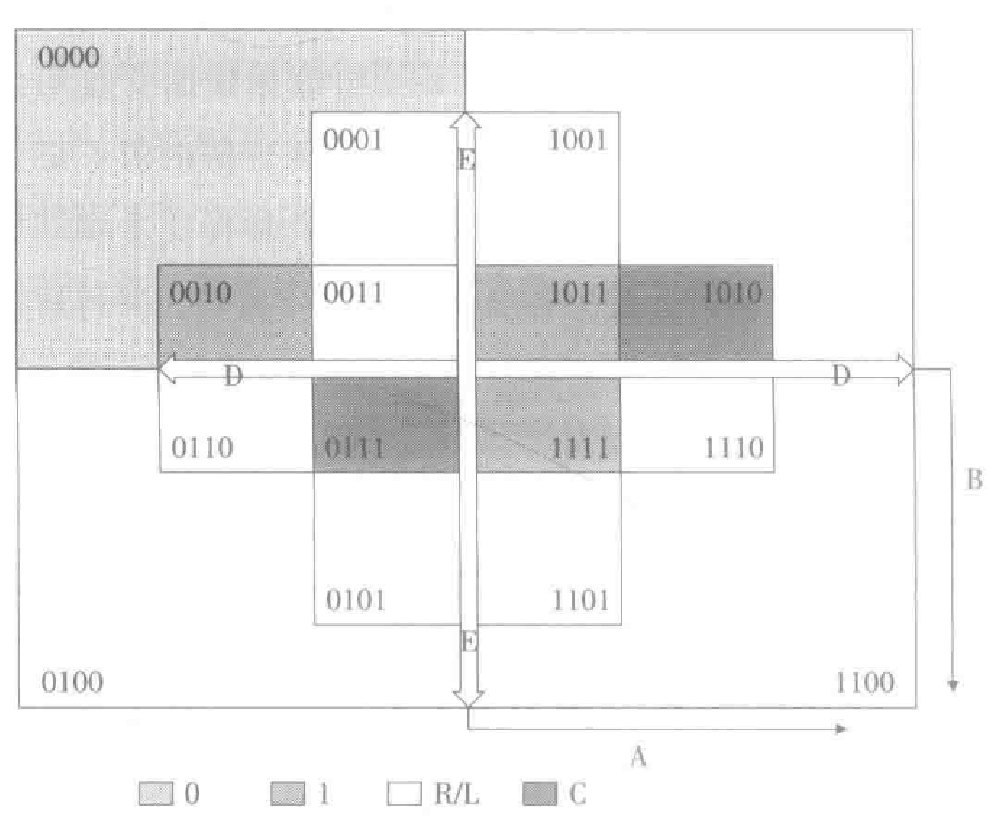
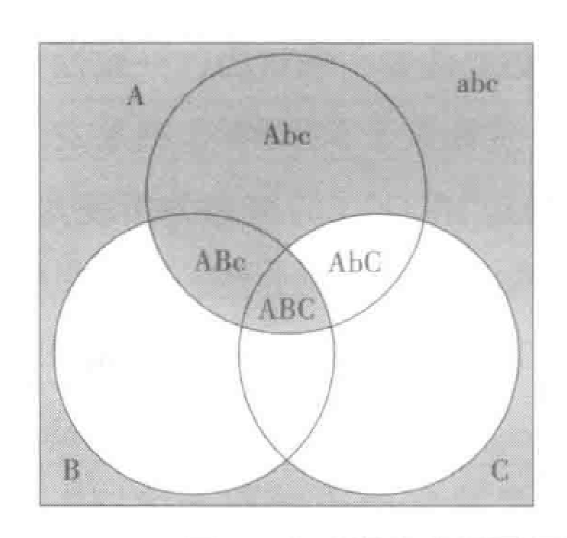


图 2 真值表可视化图(虚拟数据)

(四)处理检验与结果呈现

真值表可以直接用于QCA分析，主要包括单变量分析和条件组合分析。单变量分析的主要输出内容为各变量的一致性和覆盖

度，一致性验证单解释变量的充分必要性，一般认为，一致性大于0.8可看作结果发生的充分条件，一致性大于0.9可看作结果发生的必要条件（朱代琼、王国华，2020）。在QCA的组态分析视角中，更强调单变量的必要性和由单变量构成组态的充分性；覆盖度验证单解释变量对于结果变量的解释力度，一般包括原始覆盖度（raw coverage）、唯一覆盖度（unique coverage）和解的覆盖度（solution coverage），单变量分析结果中的覆盖度指原始覆盖度（Ragin，2000），大于0.5则认为某解释变量对结果的解释率可达50%以上。条件组合分析的主要输出有复杂方案（complex solution）、简化方案（parsimonious solution）和中间方案（intermediate solution）。复杂方案将所有变量纳入分析并直接呈现；简化方案引入部分案例样本未观察到但理论上存在的"虚拟"条件组合，反映了逻辑余数出现的情况（图3右），即简化假设（simplifying assumption）；中间方案是前两者的综合（Marx，Cambré & Rihoux，2013；凡志强，2020），优点在于不允许消除必要条件，往往是用于研究的最优选择（Dusa，2019：23-24），在不提前设定条件状态存在与否（present/absent）的情况下，复杂方案与中间方案一致（图3左）。此外，在输出结果中，覆盖度分为原始覆盖度和唯一覆盖度，原始覆盖度指该条件组合能够解释的案例占总案例的比重，唯一覆盖度指仅能被该条件组合解释的案例占总案例的比重（张晋升、祁志慧，2020），原始覆盖度大于或等于0.1，即可认为该组合具有一定的解释力（黄扬等，2019），总体一致性大于或等于0.8即证明该组合的分析有意义。



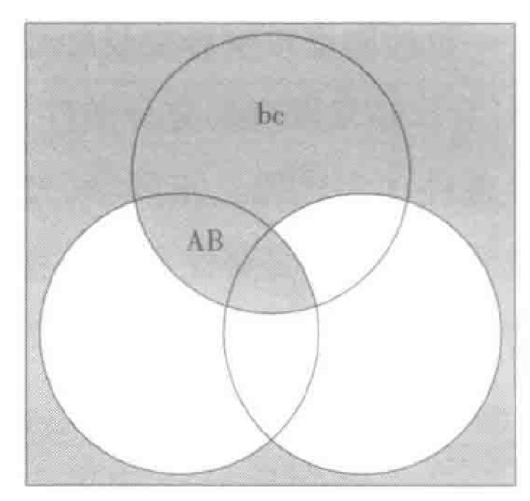


图3　条件组合分析可视化图(李蔚、何海兵,2015)

数据处理后，可以得到充分必要的单变量条件，以及对结果变量解释度较好的条件组合。目前，大多数研究的主体内容停留在这一阶段，随后直接进入结果分析或提出对策建议，但是研究者还可以根据需要对QCA结果进行稳健性（robustness）检验[①]，即检验条件组合对研究人员在过程中为降低系统复杂性所致变化的敏感程度(sensitiveness)。这一检验回应了学界对QCA不够稳定、受研究人员主观影响大的质疑。在QCA中，研究者需要尽可能降低过程的复杂性来增加结果的稳健性，其检验主要集中在三个方面：（1）与将原始数据校准为集合成员相关的阈值更改；（2）与组合相关的案例频率；（3）组合的一致性。相应地，将原始数据校准为集合成员、与组合相关的案例频率、案例一致性的可接受水平是研究者应当特别关注的三个方面（Skaaning，2011）。

①主要有两种方法支持稳健性检验：（1）三角验证（methodological triangulation），是指用多种方法论和工具进行验证，在QCA中，即将量化研究和质化研究的方法结合起来进行嵌套分析，如统计分析和过程追踪；（2）更改案例选择、模型规范、条件和结果的操作化，例如用相对狭义的案例替换更加广义的，增删理论模型中的解释因素，改变解释要素（explanans）和待解释事物（explanandum）的操作方式等。

综上，确定研究问题并选择案例样本是QCA的基础和首要步骤，其质量不仅决定了后期操作的便捷程度，更决定了数据分析时的深度和研究意义，案例选择要注重全面性与多样性，随后的变量确定与案例联系密切，更强调异质性与相关性，二者的数量主要应控制在中小规模，但具体的标准和策略尚未统一，主要依据研究者的实际需要和详细论述确定；通过确定锚值点，将原始数据校准为集合成员，并构建相应的真值表，是数据用于QCA分析的关键步骤，数据分析结果主要从单变量分析和条件组合分析的两方面情况进行解读，一致性和覆盖度分别说明了变量或相关组合的充分必要性和结果解释度。尽管目前学界多停留在数据分析阶段，但稳健性检验仍然是值得关注的部分，其构成了整个过程的完整性，有利于结果分析、问题解决、理论验证或对策提出。此外，研究者在分析时应格外注意QCA术语的使用，也可以借助图表、数字、案例等丰富表述形式，常用的有XY-坐标图（XY-plot）、维恩图（venn diagram）、真值表、关键值等。

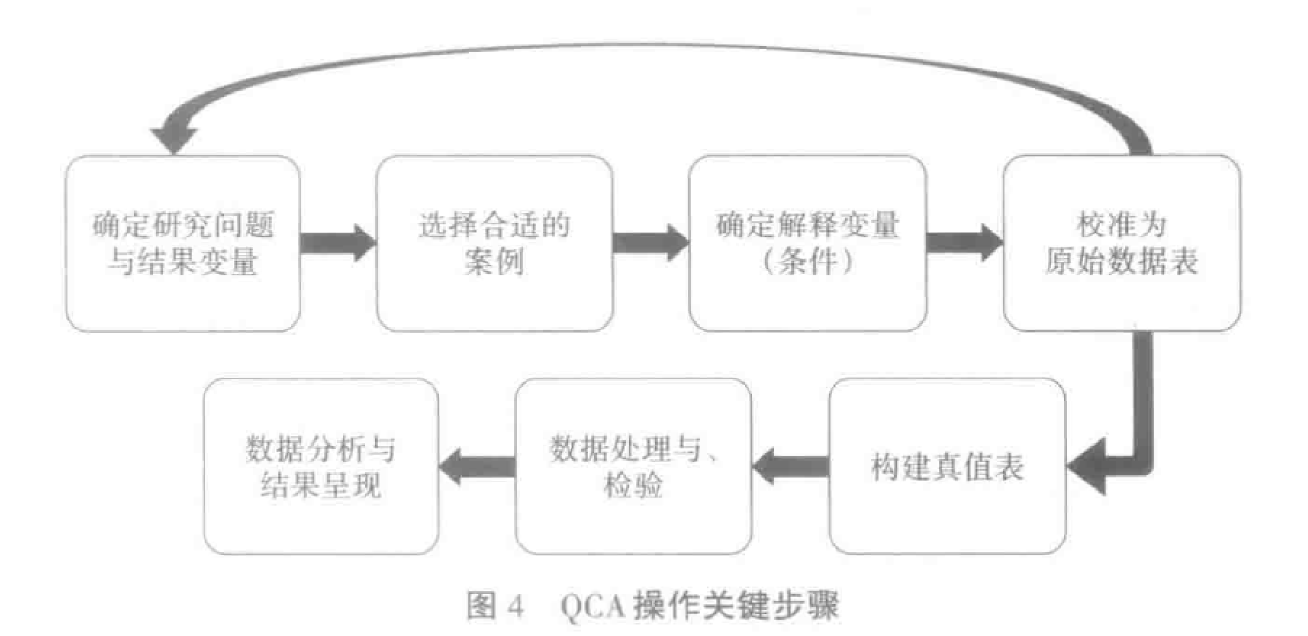


图4　QCA操作关键步骤

五、QCA的应用现状

（一）宏观应用：整体介绍与学科应用述评

QCA思想从1987年提出至今，已经有近三十年的历史，其理论内涵、操作方法和具体应用等在发展过程中逐渐明晰，引发国内外学者的广泛关注。由于QCA最早由西方学者提出，国外的研究者在前期探索上表现出更强烈的热情，体现在对QCA进行整体介绍和对相关学科研究的回顾与展望上。

拉金作为QCA的创始人，对相关思想和技术的阐释作出重要贡献，他从1987年至21世纪初先后出版了四本著作，并为fs-QCA撰写了操作手册，奠定了QCA在社会科学领域应用的基础。随后，"应用社会研究方法系列"（Applied Social Research Methods Series）出版（Rihoux & Ragin，2009），其中包括里豪克斯主编的《QCA与相关技术》（*QCA and Related Techniques*），书中不仅对拉金的思想进行了系统梳理，还介绍了cs-QCA、mv-QCA等其他工具，并从QCA原理和应用的各个部分总结了研究现状，以学界对QCA的批判和QCA的未来发展方向结尾，可以说是较早对QCA进行整体介绍的百科全书。马克思（Axel Marx）和其他学者总结了QCA前25年的发展，包括起源、最新进展和相关应用，提出QCA已经建立了自己的研究体系，正受到更多学者在更广领域的重视，强调了其对于分析多样组合因果关系的有效性（Marx et al.，2013）；还有学者从研究成果的角度切入，从数量、技术、学科、分析水平、案例与条件选择、与其他方法的结合等方面绘制了数十年间相关文献的

图景，认为QCA尽管还有诸多领域需要探索，但正逐渐成为一种主流的研究趋势（Rihoux et al.，2013）。此外，国外学者对QCA的操作过程和具体环节更为关注，倾向于发现新的研究视角，丰富理论和方法内涵。蒂姆和杜绍详细叙述了校准如何进行、一致性和覆盖度如何获得等操作过程，试图涵盖所有核心步骤，揭示QCA运行的内部原理（Thiem & Dusa，2013）；菲斯突破传统的中小规模样本概念，尝试将大规模样本创新性地引入QCA，并提出了实现的可能性（Fiss et al.，2013）；还有不少学者直接关注到了方法技术、操作关键，或提出有利于提升QCA完整性和严谨性的措施。国内学者在整体层面上的研究较少，李蔚和何海兵从QCA的基本分析逻辑、操作程序、演化与拓展、优势及局限、应用和发展趋势对其研究逻辑和应用作了整理（李蔚、何海兵，2015），但各部分过于平均，分析层面有待深入。李永发从研究设计、应用情景、关键决策点、求解过程、易错点与技巧等方面介绍了QCA的思想和内涵（李永发，2020），文章内容详实，但过于强调原理，在理解上存在一定难度。

在QCA的适用领域问题的探讨中，管理学是研究者述评较多的领域，杜运周和贾良定没有局限于QCA本身，而是首先追溯了组态理论视角和发展脉络，与传统权变视角作了区分，随后自然过渡到更加具象的QCA，讨论了其基本内容、分析过程、关键挑战等，并从管理学领域作出展望，认为QCA为管理学研究开辟了新道路（杜运周、贾良定，2017）；卡恩等学者通过分析近年来运用QCA研究管理学问题的文献，回顾了组织管理应用中的定位、策略和方向、对经济管理学研究的启示及其构型中的应用等情况，为管理学研究提供启示（Kan et al.，2016；张明、杜运周，2019；夏

鑫、何建民、刘嘉毅，2014；张驰、郑晓杰、王凤彬，2017）。在政治学领域，里豪克斯等人首次系统回顾了QCA与公共政策分析的关系，重点介绍了QCA在其中的开发应用，讨论了目前的研究成果和不足（Rihoux et al.，2011）；刘丰则提出QCA在国际关系中的研究领域、步骤技巧和方法优势、理论不足（刘丰，2015）。此外，毛湛文在将QCA引入新闻传播学科的探索上有开创性作用（毛湛文，2016）。

（二）微观应用：影响因素、路径机制、模式构型

在微观应用层面，学者运用QCA分析具体的社会问题或现象的影响因素、路径机理和模式构型三个方面，研究者选择QCA而非传统方法的原因可归纳为三点：（1）样本规模有限，难以借助统计分析工具；（2）研究现象复杂，案例本身内涵和多案例之间互动丰富；（3）因果关系隐蔽，理论上有多重因果关系组合的可能。

影响因素的研究主要是对某种现象产生或结果发生的过程分析，QCA的变量确定直接关系到具体的因素设计与验证，因研究主题不同而呈现差异。比如，在中国邻避冲突的研究上，学者发现城市业主群体更倾向于选择非暴力的抗争策略，且意见领袖作用有限（万筠、王佃利，2019）；对于突发事件中网民社会情绪产生的影响因素，学者总结为利益诉求、事件类型、规则失守、安全危机、利益抗争（朱代琼、王国华，2020）；平台企业的员工道德水平、内部活动控制力、评价标准、外部监督为其社会责任缺失提供前因构型，引导现实行为改善（彭本红、葛娇娇，2022）。可见，影响因素的相关研究主题范围广泛、研究数量丰富并形成了一定的范式，其结论对剖析具体现象或实现理想结果具有意义。

路径机理的研究多与影响因素关联，但相比于影响因素的系统性和操作性更强，现有研究多在分析影响因素的基础上，进一步归纳出路径机理，充分利用QCA的组态分析特征。以创业生态系统理论分析典型双创示范县，驱动乡村高产业和非高产业振兴质量的路径分别被归纳为要素聚合型、科技驱动型、结构优化型，以及资金抑制型和智力抑制型（郝政等，2022）；李月和曹海军紧密结合时代热点，从健康码分析省级政府数字治理的影响因素和实施路径，得出能力主导型、需求主导型与任务主导型三种路径（李月、曹海军，2020）；杨志和魏姝在研究政策爆发机理时，由对影响因素的归纳分析，提出以可操作性、中央密集干预和整体性推进策略组成的简约路径，以及该路径与不同层级政府决策注意力的嬗变耦合机制会加速政策爆发（杨志、魏姝，2020）；公共危机多元主体协同治理的七个影响因素与四条路径也密切相关（山少男、段霞，2022）。路径机理的相关研究以影响因素确定为基础，同时又是对影响因素的深化和系统化，其结论将理论上的做法具体化到实践中，更易理解和操作。

模式构型的研究较前两方面研究对总结和提炼的要求更高，多对组合情况再次进行解读分析，普适性更强，但有时也将路径和模式等同，研究者利用QCA呈现复杂因素的内部情况，根据推导出的多重解释组合建立模型。在政策转移中，政策再生产呈现三种模式：资源主导型，行政压力、资源、社会力量复合驱动型，政府和智库合作型（熊烨、周建国，2017）；省级政府数据开放高水平绩效的生成模式则包括组织驱动型、技术驱动型、经济-技术驱动型和综合驱动型四种（汤志伟、罗意，2021）；在组态视角下，我国邻避冲突产生模式可概括为利益风险失衡模式、决策信任制约模

式、技术条件助推模式（王洛忠、孙枭坤、陈宇，2020）。模式构型的相关研究主要根据各因素的显著性不同进行提炼总结，其结论是对数据分析结果的再次整理和解读，使全文结构更加符合QCA的组态思想和特征。

综上，国内学者更习惯直接将QCA当作研究工具和分析方法，遵循QCA的操作范式，解读具体的社会现象或提出对策，对于QCA的实际应用具有丰富和拓展的作用，但目前的研究主题仍然具有局限性，且在变量设置上创新性不足，相较国外仍有较大发展空间。

六、总结和讨论

定性比较分析（QCA）是一类研究思想和分析技术的总称。作为研究思想，其代表了集合论和布尔代数体现的组态视角，意味着将每个案例都视作整体，通过案例间的对比分析，得到不同条件和条件组合对结果的解释程度；作为分析技术，其代表了cs-QCA、fs-QCA、mv-QCA、T-QCA等多样的方法工具。在具体操作中，研究者需要根据研究内容选择案例、确定变量、进行校准和构建真值表，分别对单变量和条件组合进行处理分析，并可以检验稳健性，最终以多样形式呈现结果。从应用现状来看，国外学者更注重宏观层面的介绍和创新，而国内学者更善于微观层面的现象分析和问题解决。

定性比较分析对于处理因果复杂性、精简案例变量、得出组合解释等具有优势，平衡了定性研究的深度与定量研究的广度，对其的尝试实际上是顺应了多元方法、融合研究的趋势，学界已有越来

越多的学者进一步提倡QCA与其他方法的结合（Creswell，2009：205-207）。同时，应该注意到QCA存在的一些局限性，主要表现在：（1）案例敏感性（case sensitivity）过强，QCA是基于案例的分析，在过程中将每个案例看作一个复杂整体，分析其内部的影响变量和组合，在数据处理时，增删某个或某些案例都会使条件组合和结论发生变化，甚至会产生矛盾，但矛盾并非指向研究失败，如前文所述，也可能会推动研究者在理论、案例、条件等方面进行反思与完善；（2）案例和条件数量有限制，主流的QCA分析案例和条件都使用中小规模样本，不利于结论的普适性，而布尔代数的运算法则为N个条件变量产生2^N个条件组合，个别数量的变动都可能带来几何级别的效应，虽然已有学者提出将QCA引入大样本研究，但尚没有具体应用；（3）变量校准的划分标准不够客观，最初的QCA分析仅使用cs-QCA，二元阈值的设定局限了变量划分，同时，变量的确定主要是基于理论研究和实践经验，并没有针对某一研究问题的统一标准，增强了其中的主观性和不确定性，尽管后来增加了fs-QCA、mv-QCA等技术，但仍不能很好解决这一问题；（4）时序问题暂不能得到有效解决，目前，用于QCA研究的主要方法和数据分析都以时序性不会影响结果为假设前提，但并不能因此完全排除时序的作用，T-QCA和TS-QCA等技术的提出意味着学界已经注意到这一问题，但相关思想仍然不够成熟，不能从根本解决时序问题；（5）QCA的因果解释不够全面，虽然是超越定性研究与定量研究的"第三条道路"，但并不能完全取代传统的相关分析、回归分析等，其结果的解释力和有效性存在争议。

QCA的局限性警示了研究过程中的一些关键环节：研究者必须充分重视有意义的矛盾组态和逻辑余数，并不断在案例和理论间进

行“对话”，以减弱案例敏感性的不利影响；应有机结合定量研究方法与定性研究方法的优势，既保持程序客观性与科学性，严格进行案例编码与运算，又不损失案例独特性，在深度阐释个案的基础上，明确具有相似性的案例集群。此外，研究者可以有针对性地尝试探索QCA对大规模样本和时序问题的适用性，或考虑跳出QCA多种技术的方法孤岛，如以案例分析验证和解释路径组合（叶成城，2021）、以文本挖掘充实研究对象和以回归分析辅助结果解读（朱逸、桂勇，2020），但目前的方法尝试仍围绕QCA展开，如何使之与传统方法构成优势互补的混合研究法尚待探索。

总体来说，QCA作为一种新的研究视角，在近几年受到国内学者的广泛关注，从研究数量和研究学科上来看，目前均保持稳步发展的趋势。笔者认为，QCA的优势与不足并存，机遇与挑战同在，了解其理论和方法内涵、操作步骤与策略、当前的应用情况等有利于培养新思维和作出新尝试，但QCA仍伴随诸多不足与质疑，其方法完善和应用前景等仍有待在未来研究中进一步加强。

参考文献

陈宇、闫倩倩，2019，《“中国式”政策试点结果差异的影响因素研究——基于30个案例的多值定性比较分析》，《北京社会科学》第6期：42-52。

程聪、贾良定，2016，《我国企业跨国并购驱动机制研究——基于清晰集的定性比较分析》，《南开管理评论》第6期：113-21。

程代展，赵寅，徐相如，2011，《从布尔代数到布尔微积分》，《控制理论与应用》第10期：1513-23。

迟永，2014，《美国介入领土争端的行为——基于模糊集定性比较分析的解释》，《世界经济与政治》第10期：56-80+156-157。

杜运周、贾良定，2017，《组态视角与定性比较分析（QCA）：管理学研究的一条新道路》，《管理世界》第6期：155-67。

凡志强，2020，《垃圾分类政策的扩散机制与效果研究——基于42个试点城市的定性比较分析》，《地方治理研究》第4期：60-76+79。

郝政、何刚、王新媛、张勇，2022，《创业生态系统组态效应对乡村产业振兴质量的影响路径——基于模糊集定性比较分析》，《科学学与科学技术管理》第1期：57-75。

黄荣贵、桂勇，2009，《互联网与业主集体抗争：一项基于定性比较分析方法的研究》，《社会学研究》第5期：29-56+243。

黄荣贵、郑雯、桂勇，2015，《多渠道强干预、框架与抗争结果——对40个拆迁抗争案例的模糊集定性比较分析》，《社会学研究》第5期：90-114+244。

黄扬、李伟权、郭雄腾、段晶晶、曹嘉婧，2019，《事件属性、注意力与网络时代的政策议程设置——基于40起网络焦点事件的定性比较分析（QCA）》，《情报杂志》第2期：123-30。

李蔚、何海兵，2015，《定性比较分析方法的研究逻辑及其应用》，《上海行政学院学报》第5期：92-100.

李永发，2020，《定性比较分析：融合定性与定量思维的组态比较方法》，《广西师范大学学报（哲学社会科学版）》第3期：89-110。

李月、曹海军，2020，《省级政府数字治理影响因素与实施路径——基于30省健康码应用的定性比较分析》，《电子政务》第10期：39-48。

里豪克斯·伯努瓦、拉金·查尔斯，2017，《QCA设计原理与应用》，北京：机械工业出版社。

刘丰，2015，《定性比较分析与国际关系研究》，《世界经济与政治》第1期：90-110+158-159。

刘晓亮、侯凯悦、张洺硕，2019，《从地方探索到中央推广：政府创新扩散的影响机制——基于36个案例的清晰集定性比较分析》，《公共管理学报》第3期：157-167+176。

毛湛文，2016，《定性比较分析（QCA）与新闻传播学研究》，《国际新闻界》第4期：6-25。

彭本红、葛娇娇，2022，《基于QCA方法的平台企业社会责任缺失的影响因素分析》，《软科学》第1期：69-76。

山少男、段霞，2022，《复杂性视角下公共危机多元主体协同治理行为的影响

因素与行动路径——基于元分析与模糊集QCA的双重分析》,《公共管理与政策评论》第1期:104-19。

释启鹏,2016,《时间中的定性比较分析:TQCA与TSQCA的发展》,《比较政治学研究》第1期:40-58。

汤志伟、罗意,2021,《资源基础视角下省级政府数据开放绩效生成逻辑及模式——基于16省数据的模糊集定性比较分析》,《情报杂志》第1期:157-64.

万筠、王佃利,2019,《中国邻避冲突结果的影响因素研究——基于40个案例的模糊集定性比较分析》,《公共管理学报》第1期:66-76+172。

王洛忠、崔露心,2020,《公民参与政策制定程度差异的影响因素与路径模式——基于31个案例的多值定性比较分析》,《南京大学学报(哲学·人文科学·社会科学)》第6期:99-111+159-160。

王洛忠、孙枭坤、陈宇,2020,《组态视角下我国邻避冲突产生模式概化——基于30个案例的定性比较分析》,《城市问题》第6期:47-55。

夏鑫、何建民、刘嘉毅,2014,《定性比较分析的研究逻辑——兼论其对经济管理学研究的启示》,《财经研究》第10期:97-107。

熊烨、周建国,2017,《政策转移中的政策再生产:影响因素与模式概化——基于江苏省“河长制”的QCA分析》,《甘肃行政学院学报》第1期:37-47+126-127。

杨志、魏姝,2020,《政策爆发生成机理:影响因素、组合路径及耦合机制——基于25个案例的定性比较分析》,《公共管理学报》第2期:14-26+165。

叶成城,2021,《第二波现代化之“帝国的共鸣”——1848年欧洲变革的案例研究、定性比较分析》,《欧洲研究》第4期:62-88+6-7。

张驰、郑晓杰,王凤彬,2017,《定性比较分析法在管理学构型研究中的应用:述评与展望》,《外国经济与管理》第4期:68-83。

张晋升,祁志慧,2020,《虚假新闻高热度传播组合路径研究——基于57例虚假新闻的清晰集定性比较分析(QCA)》,《西南民族大学学报(人文社科版)》第2期:178-87。

张明、杜运周,2019,《组织与管理研究中QCA方法的应用:定位、策略、方向》,《管理学报》第9期:1312-23。

周俊、王敏,2016,《网络流行语传播的微观影响机制研究——基于12例公共事件的清晰集定性比较分析》,《国际新闻界》第4期:26-46。

朱代琼、王国华，2020，《突发事件中网民社会情绪产生的影响因素及机理——基于三元交互决定论的多个案定性比较分析（QCA）》，《情报杂志》第3期：95–104。

朱逸、桂勇，2020，《网络直播购物：影响消费者行动参与的信息策略选择——基于“文本挖掘+QCA”的混合性研究》，《企业经济》第5期：95–103。

Amenta, Edwin & Jane D Poulsen 1994, “Where to Begin: A Survey of Five Approaches to Selecting Independent Variables for Qualitative Comparative Analysis.” *Sociological Methods & Research*23(1):22–53.

Basurto, Xavier & Johanna Speer 2012, “Structuring the Calibration of Qualitative Data as Sets for Qualitative Comparative Analysis(QCA).”*Field Methods* 24(2):155–74.

Caren, Neal & Aaron Panofsky 2005, “TQCA: A Technique for Adding Temporality to Qualitative Comparative Analysis.” *Sociological Methods & Research* 34 (2): 147–72.

Creswell, John W. 2009, *Research design: qualitative, quantitative, and mixed methods approaches* Thousand Oaks, Calif: Sage Publications.

Cronqvist, Lasse 2004, “Presentation of TOSMANA Adding Multi–Value Variables and Visual Aids to QCA.”*COMPASSS Working Papers*2004(20).

Fiss, Peer C. 2007, “A Set–Theoretic Approach to Organizational Configurations.” *Academy of Management Review*32(4):1180–98.

Fiss, Peer C. 2011, “Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research.” *Academy of Management Journal* 54 (2): 393–420.

Fiss, Peer C., Dmitry Sharapov & Lasse Cronqvist 2013, “Opposites Attract? Opportunities and Challenges for Integrating Large–N QCA and Econometric Analysis.”*Political Research Quarterly*66(1):191–98.

Gerring, John 2007, *Case Study Research: Principles And Practices*. Cambridge: Cambridge University Press.

Grofman, Bernard & Carsten Q. Schneider 2009, “An Introduction to Crisp Set QCA, with a Comparison to Binary Logistic Regression.”*Political Research Quarterly* 62 (4):662–72.

Hino, Airo 2009, "Time-Series QCA Studying Temporal Change through Boolean Analysis." *Sociological Theory and Methods*24:247-65.

Hume, David 1999, *An enquiry concerning human understanding*. Oxford; New York: Oxford University Press.

Longest, Kyle C. & Stephen Vaisey 2008, "Fuzzy: A Program for Performing Qualitative Comparative Analyses (QCA) in Stata." *The Stata Journal: Promoting Communications on Statistics and Stata*8(1):79-104.

Marx, Axel, Bart Cambré & Benoît Rihoux 2013, "Crisp-Set Qualitative Comparative Analysis in Organizational Studies." In P C Fiss, B Cambré & A Marx (ed.), *Configurational Theory and Methods in Organizational Research (38), Research in the Sociology of Organizations*. Emerald Group Publishing Limited.

Marx, Axel & Adrian Dusa 2011, "Crisp-Set Qualitative Comparative Analysis (CsQCA), Contradictions and Consistency Benchmarks for Model Specification." *Methodological Innovations Online*6(2):103-48.

Marx, Axel, Benoît Rihoux & Charles Ragin 2013, "The Origins, Development, and Application of Qualitative Comparative Analysis: The First 25 Years." *European Political Science Review* 6(1):115-42.

Meyer, Alan D., Anne S. Tsui & C. R. Hinings 1993, "Configurational Approaches to Organizational Analysis." *The Academy of Management Journal*36(6):1175-95.

Mill, John Stuart 2002, *A System of Logic: Ratiocinative and Inductive*. Honolulu: University Press of the Pacific.

Mozas-Moral, Adoración, Encarnación Moral-Pajares, Miguel Jesús Medina-Viruel & Enrique Bernal-Jurado 2016, "Manager's Educational Background and ICT Use as Antecedents of Export Decisions: A Crisp Set QCA Analysis." *Journal of Business Research* 69(4):1333-35.

Ordanini, Andrea, A Parsu Parasuraman & Gaia Rubera 2013, "When the Recipe Is More Important Than the Ingredients." *Journal of Service Research* 17:134-49.

Przeworski, Adam & Henry Teune 1970 *The logic of comparative social inquiry*. New York: Wiley-Interscience.

Ragin, Charles C. 1987 *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California press.

Ragin, Charles C. 2000 *Fuzzy-Set Social Science*. Chicago: University of Chicago

Press.

Rihoux, Benoît 2006, "Qualitative Comparative Analysis(QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research." *International Sociology* 21(5):679-706.

Rihoux, Benoît, Priscilla Álamos-Concha, Damien Bol, Axel Marx & Ilona Rezsöhazy 2013, "From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011." *Political Research Quarterly* 66(1):9-18.

Rihoux, Benoît & Bojana Lobe 2009. "The Case for Qualitative Comparative Analysis (QCA): Adding Leverage for Thick Cross-Case Comparison." Pp. 222-42, In David Byrne & Charles Ragin, *The SAGE Handbook of Case-Based Methods*. London: SAGE Publications Ltd.

Rihoux, Benoît & Charles Ragin C. 2009 *Configurational comparative methods: qualitative comparative analysis (QCA) and related techniques*. Thousand Oaks: Sage.

Rihoux, Benoît, Ilona Rezsöhazy & Damien Bol 2011, "Qualitative Comparative Analysis (QCA) in Public Policy Analysis: An Extensive Review." *German Policy Studies* 7(3):9-82.

Schneider, Carsten Q. & Claudius Wagemann 2006, "Reducing Complexity in Qualitative Comparative Analysis (QCA): Remote and Proximate Factors and the Consolidation of Democracy." *European Journal of Political Research* 45(5):751-86.

Schneider, Carsten & Claudius Wagemann 2010, "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." *Comparative Sociology* 9:397-418.

Seny Kan, Anderson Konan, Emmanuel Adegbite, Sami El Omari & Mahamat Abdellatif 2016, "On the Use of Qualitative Comparative Analysis in Management." *Journal of Business Research* 69(4):1458-63.

Skaaning, Svend-Erik 2011, "Assessing the Robustness of Crisp-Set and Fuzzy-Set QCA Results." *Sociological Methods & Research* 40:391-408.

Thiem, Alrik & Adrian Duşa 2013, "QCA: A Package for Qualitative Comparative Analysis." *The R Journal* 5(1):87-97.

Vink, Maarten P. & Olaf Van Vliet 2009, "Not Quite Crisp, Not Yet Fuzzy? Assessing the Potentials and Pitfalls of Multi-Value QCA." *Field Methods* 21(3):265-89.

王路．学术研究：瞄准无知的认知行动——读默顿和斯梅尔瑟的学术回忆录[M/OL]//赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

学术研究：瞄准无知的认知行动
——读默顿和斯梅尔瑟的学术回忆录

王路[①]

摘要：默顿（Robert K. Merton）与斯梅尔瑟（Nell J. Smelser）分别在1987年、1969年发表回忆录，从现象、理论、经验三方面讨论如何寻找社会学研究的问题，它们对青年研究者如何选择自己的学术研究主题和问题，具有启发作用。具体来说，青年研究者可通过发现社会学知识库中的空白或无知，并设定这种无知的边界、关注它和已知理论和事实的联系，避免理论与经验的割裂，在起步阶段就树立正确的研究方向：由详解的无知走向理论发展和经验研究。

关键词：详解的无知；研究问题；理论；事实

Abstracts: Robert K. Merton and Nell J. Smelser published academic memoirs in 1987 and 1969, respectively discussing how to find sociological research in three aspects: phenomena, theories, and materials.

①作者简介：王路，中国社会科学院大学社会与民族学院博士研究生，联系方式：mucwanglu2016@163.com。

They are instructive for young researchers in how to choose their academic research problems and questions. Specifically, young researchers can avoid the fragmentation of theory and experience by identifying specified ignorance in the sociological knowledge stock, setting the boundaries of this ignorance, paying attention to its connection with known theories and facts, and setting the right research direction at the beginning: from specified ignorance to theory development and empirical research.

Key words: Specified Ignorance; Research Problems; Theories; Facts

一、学术研究的定位

学术研究的起点与核心在于有问题意识，问题意识的明确是一个从模糊到清晰的过程，需要经验与理论两方面加力：在经验层面，它源于研究者对于社会事实的“困惑”和“疑问”，所呈现的是无法触及问题核心的、模糊的“表现”；在理论层面，已有研究的脉络与倾向可以为研究者提供一些理解、探究“表现”的方向，在特定的方向指引中，研究者可以逐步明晰“表现”的一些核心特征、关键变量。据此，我们可以满足涂尔干对研究者所提出的基本要求，即能够让自己和别人知道自己在研究什么。而问题在于，尤其对于青年研究者而言，学术研究的第一步往往也是最难的一步，“没有问题意识”“理论与经验之间两层皮”“研究对象不明确”等问题频繁出现在各类学术研究中。因此，如何探索研究的可能、确定研究问题是青年研究者在学术道路中的一项重要课题。

对于以上问题，我们可以从默顿（Merton，1987）和斯梅尔瑟

（Smelser，1969）的学术回忆录中找到一些提示：斯梅尔瑟回溯个人学术生涯的主线是自己如何确定研究问题并围绕其中未回应的问题继续推动研究，默顿则从确定待研究现象（establishing the phenomenon）、详解的无知（specified ignorance）、战略性的研究材料（strategic research materials，SRMs）三个方面讨论科学研究中的实践活动。在详细论述中，默顿和斯梅尔瑟都谈及现象、理论与经验是如何带动自己持续开展学术研究的，而对于青年读者而言，以上内容则启发我们思考在学术研究中的两个核心问题：在学术生涯的起步阶段，如何确定学术研究的起点？又如何以此持续推进学术研究？对此，青年研究者可以从默顿所提出的"详解的无知"（Merton，1987）出发，通过识别已有研究中的无知，把握详解的无知与学术研究之间的关系，重新审视理论与经验之间的关系，寻找研究的可能，推进理论与经验研究的发展。

二、识别详解的无知

默顿用"详解的无知"提醒研究者重视已有研究中的无知或空白，要有识别已有研究中无知的意识和能力，唯有此才能更好地推进研究。斯梅尔瑟的研究经历则进一步为理解"详解的无知"提供了经验案例，他详细描述了自己如何基于已有研究中未解答的问题不断推进他对于集体行动的研究。如果没有默顿的提示，我们或许会感叹斯梅尔瑟的幸运，他可以确定他的研究主题并不断扩展研究问题，逐步构建出自己学术生涯的研究体系。然而，斯梅尔瑟能够幸运地构建自己的研究体系需要有两个基本前提：首先，他善于感知已有研究中所未回应的或忽略的问题，即具有识别"详解的无

知”的意识与能力；其次，他认真投入到回应“详解的无知”的行动中，始终以已有的无知作为当下研究的起点，继续构建理论与事实的关联，推进研究的发展。

对于刚进入某一研究领域的青年研究者而言，我们极易混淆在已有研究所构建的有限知识库与不断延展的经验材料之间的矛盾关系中，善于理解、把握文献知识库中的已知，而相对忽视其中的未知。在具体研究中，研究者如果只需要借助已掌握的理论或概念就足以呈现其所观察到的所有经验材料，那这样的学术研究从一开始就会很顺利，而且随着理论和概念的累积，会越容易发现与理论、概念相适的经验事实。而事实上，理论或概念与经验材料之间的矛盾关系时刻存在于社会科学研究中，也因此产生了青年研究者对科学研究的一些误解，比如会将已有理论奉为“想当然”的圭臬，以此检视经验材料中的问题，或用与众不同的经验材料来验证已知理论的正确性。当然，这种以理论定义经验材料的策略非常有用，它可以积累经验材料、增加关键变量来发展理论，但也有研究的代价，以科层制为例，我们通常直接将它作为分析经验材料的概念工具，比如自上而下的权责体系、特定的规则与程序等，当发现与概念所包含的某些特质相匹配的经验材料，便自信地得出结论“这就是科层制”，而忽视了科层制与经验材料之间的距离，它不是一种简单的经验归纳，更适合被描述为一种经验的概推，以至于我们看不到在先前的概推过程与当下的经验研究中所舍弃的可能更有意思的研究问题（贝克尔，2017：20），也就无法更好地回答“科层制究竟是什么”等问题。更何况，我们对理论的忽视可能是多方面的，相应地，被忽视的经验材料会始终存在、发生变化并影响着研究对象。以此，理论与经验反而会在科学研究的发展中逐步构建了一种无知的边界，

加剧理论与经验之间的割裂。

从研究的过程来看，无知的边界割裂的是理论“我”与经验“我”的整体性关系，它所体现的是经验“我”无法感知理论中的缺失，理论“我”亦无法引导经验“我”去追问现实中的困惑。研究结论完全取决于研究者使用怎样的理论假定，唯一的好处是能够让经验“我”很容易找到“应该被发现”的“表现”，让看到的“表现”越来越符合理论“我”选择的概念，结果确是因无法了解真实的经验常态而加剧理论与经验之间的割裂：理论“我”似乎充分掌握了研究的“套路”，从理论出发，“切割”经验，经过一番艰苦、复杂的推论，得出一个理论的常识，完成了“理论-经验-理论”的“抽象大循环”，只是在这个大循环中，理论“我”忘记了持续进行自我批判性的思考（埃尔斯特，2019：34），也很难体会到经验质感的研究意涵；经验“我”则从经验出发，四处搜寻可能与经验材料相关的概念名词，用经验拼凑概念，既无法通过社会学研究方法看到与经验材料相关的潜在可能，也无法通过理论去思考经验材料中所呈现的差异，更不能将经验现象所反映的问题放到已有的研究脉络中去审视，最终只是没有理论视野的经验“我”讲了一个没有问题意识的故事，以此形成囿于经验材料而无理论关怀的“经验小循环”。

以上两种循环都无法真正实现理论与经验之间的对话，哪怕在不同的研究中能看到理论或经验的身影，也不过是理论“切割”经验或经验“硬套”理论的不同表现形式而已（马丁，2020：25）。我们善于将既有理论纳入已知的知识库存中，也善于借助理论去寻找经验材料中较为笼统的共识，只是不善于通过学术研究去发现更多隐藏的、被忽视的却一直存在的经验事实。至此，我们或许能够更

深刻地理解斯梅尔瑟的研究感悟，没有一种解决方案能够视为对研究问题的最终解答，尚未解答的问题始终存在，研究成果所带来的巨大满足感是相对的，只有遗留的问题才是推进研究的恒久动力（Smelser，1969：158）。

三、瞄准无知的认知行动

默顿从功能视角认识“详解的无知”，以此关联学术研究的起点与过程。青年研究者不仅要有识别无知的意识，更要有应对无知的行动，即能够基于“详解的无知”提出一个具有创新性的、有价值的、有待解决的研究问题，以此推进理论发展与经验研究。在回忆录中，默顿以越轨行为研究为例，呈现该研究如何在行为学习论、社会结构论、标签理论、冲突理论四种思想传统中得以发展；斯梅尔瑟结合个人研究历程回溯集体行为研究的推进过程，当他发现最近的研究无法解释个体心理压力与集体行动之间的关系时，在1961—1968年，他的主要工作便是从社会学和心理学的角度继续构建他的研究对象。这样看来，研究可能性始终伴随研究成果而出现，只是其内在机理不单是对学术研究中的已知的积累，而是基于“详解的无知”的探索行动。

因此，从学术研究过程而言，研究可能与“详解的无知”之间存在一种辩证发展关系：学术研究中的知识库存越丰富，研究者对于无知的感知与困惑也会越多，并可以基于无知的困惑而探索更多的研究可能。对此，我们可以尝试把自己视为理论发展与经验研究的观察者。对于理论而言，任何一种理论的提出与发展都是研究者对“详解的无知”主动进行反问、求证的结果，研究者可以以理论

为切点，把握分析经验材料的脉络，并不断接受经验材料的补充与质疑，寻找发展理论的可能；经验研究亦如此，面对经验材料，研究者需要借助理论、研究方法的引导去观察之前所未观察到的社会事实，发现更多不易识别的特征，在始终无法脱离复杂性特征的观察与描述中，逐步提出对经验材料的新思考，以确保经验研究的理论贡献。从这一层面来看，已有研究中的无知虽然证实了研究成果的不完美，但也以此赋予了研究成果更持久的生命力。正如韦伯所提出的"理想类型"（Weber，1949/1904：90），它的生命力正是在于其片面性，它本身是舍弃了很多实在特征的指引工具，而不是一个能够包含实在所有特征的全能模型，我们可以把它理解为一个帮助研究者寻找可能性的工具，可以指引研究者做归因判断，帮助研究者超越已有的理想类型、发现新的理想类型。

如此，我们便能进一步理解斯梅尔瑟对"详解的无知"的坚守，面对已有研究，始终要怀有对研究结果的"不满"之心去观察、思考一番，寻找尚未解答的问题，明晰具有创新性的研究可能，结合理论、经验推动具体研究。例如，威廉·富特·华特（Whyte，1943）在《街角社会》一书中所回应的是当时贫民窟研究中的"详解的无知"，他通过对意大利人贫民区所做的实地研究，得出与芝加哥学派不同的结论，认为贫民窟具有自身的组织原则，而并非当时学界所定义的非组织化社区。

至此，我们或许会提醒自己应该努力避免学术研究中的唯名论倾向，对于经验事实的观察与描述不能取决于理论，理论的功能是引导而非限定。更进一步说，我们需要借助已有研究中的"详解的无知"，推动理论发展与经验研究。学术研究就好像是建造一座从知识通向无知的未完工的桥，我们不能仅作为欣赏者、借鉴者去感

叹、模仿其他建桥人的精湛技艺，而忘了自己也是建桥人，面对这座桥，我们必须要知道别人已经把桥修到哪里了，过程中遇到了哪些问题，不同人对这一问题持有怎样不同的看法，解决了哪些方面，未解决哪些方面，我们又能对这个问题有怎样的理解，又如何能在已有研究的基础上提出自己的研究方案。这种基于详解的无知的认知行动，能够让青年研究者在学术研究中同时兼顾两个重要方面：一面是深入到学科的根基，体悟理论的厚重，基于社会学于发展中所形成的学科传统去探寻社会在时空中的形态与结构；另一面则要外延至学术研究的边缘处，接受经验事实的混乱或偏离，探索现代社会的结构性特质，以拓展学科的研究可能。

四、总结与反思

默顿和斯梅尔瑟在回忆录中讨论了他们对社会学研究中的现象、理论与经验的看法，其中，“详解的无知”是理解三者对于学术研究意义的核心。青年研究者可以通过“详解的无知”确定研究问题，反思当下研究中理论与经验之间的断裂，并可以尝试以观察者的角色在理论与经验之间不断反问、求证，以此寻找更多的研究可能。在此过程中，青年研究者不仅是观察者，亦是行动者，“详解的无知”促使研究者既要能深触到学科的理论根基，又要能拓展研究的经验外延，以此推动理论发展与经验研究。

当然，我们也要注意到默顿和斯梅尔瑟是从方法的角度启发读者反思理论与经验在研究中的关系与功能，在具体研究中，理论与经验之间的整体关系需要围绕研究对象来呈现，这一点可以从斯梅尔瑟的学术经历中看到他如何围绕集体行动不断拓宽自己的研究维

度。对于任何一项研究而言，研究者既要增强理论视野的敏感力，也要加强经验视野的洞察力，而这两者的核心指向应是研究对象的构建，只有围绕研究对象进行多维度的问题思考，才能逐步加深其对研究对象的想象与把握，进而推进理论发展与经验研究。

参考文献

霍华德·贝克尔，2017，《社会学家的窍门：当你做研究时你应该想些什么?》，陈振铎译，重庆：重庆大学出版社。

乔恩·埃尔斯特，2019，《解释社会行为：社会科学的机制视角》，刘骥、何淑静、熊彩等译，重庆：重庆大学出版社。

约翰·李维·马丁，2020，《领悟方法：社会科学研究中的方法误用及解决之道》，高勇译，重庆：重庆大学出版社。

Smelser, Nell J. 1969,"Some Personal Thoughts on the Pursuit of Sociological Problems."*Sociological Inquiry* 39(2).

Merton, Robert K. 1987,"Sociologist's Notebooks: Establishing the Phenomenon, Specified Ignorance, and Strategic Research Materials."*Annual Review of Sociology* 13(1).

Weber, Max 1949, *The methodology of the social sciences*. New York, NY: Free Press.

Whyte, William Foote 1943, *Street corner society: The social structure of an Italian slum*. Chicago, IL: University of Chicago Press.

魏廷烨．打磨田野手艺：北京市 A 区日结工市场调[M/OL]//赵联飞，赵锋．社会研究方法评论：第 2 卷．重庆：重庆大学出版社．

打磨田野手艺：北京市A区日结工市场调查反思

魏廷烨[①]

摘要：田野调查是一门需要不断打磨的手艺。本文通过对既往所做北京市A区日结工市场调查的复盘，来检讨调查得失和打磨田野手艺。以寻工磋商为主线，整个田野资料的整理与组织形成了市场结构、市场规模、寻工磋商三大板块。前两大板块共同形成了寻工磋商的结构性限制，而寻工磋商则为前两大板块增添了变数与能动。三者共同构成了日结工市场的完整断面描述。随后，在复盘调查所遇理论困境中，笔者着重强调了描述性研究工作的基础性与重要性。这也是本文舍去繁杂解释性理论并重新组织调查资料的初衷。

关键词：市场结构；日结工作；寻工磋商；描述性研究

Abstract: Fieldwork is a craft that should be persistently sharpened up. By reviewing the investigation on day labor market in district A of

①作者简介：魏廷烨，中国社会科学院大学社会与民族学院本科生，联系方式：307887215@qq.com。

Beijing, this paper tries to reflect on the gains and losses of the study so as to strengthen the craft of fieldwork. Threaded by job searching and matching, all materials can be divided into three parts: market structure, market size, job search and match. The first two parts form the structural constraints of job search and match, but the third part also adds contingencies and initiatives to them. They horizontally describe the day labor market as a whole. Then, after rethinking the theoretical dilemmas encountered in the investigation, it strikes me how basic and important the descriptive research is. This is the reason why I have to cut down lots of explanatory theories and reorganize the materials.

Key words: Market Structure; Day Labor Work; Job search and Match; Descriptive Research

一、前言

本文并不试图通过分解、剖析和凝练田野资料以求关于因果机制的解释理论，也并不试图构建抽象的理论模型。即使这种理论模型能够作为田野调查中的种种现象的参照系，观照出不同类别和层次的繁多意义。本文所试图做的，只是通过整理和组织田野资料来复盘整个调查过程，并从复盘中检讨调查得失，以求为其他研究者和笔者自己今后的田野调查予以警醒与反思。因此放弃学术论文写作的八股样式，删去冗余的种种理论与经验文献，仅以田野材料的组织和反思供读者审读。

本次田野调查的对象是北京市A区日结工市场，调查方式是参与观察与半结构式访谈。该市场脱胎于A区大型建材市场附近的马

路劳工市场。马路边劳工市场是一个大城市内的现实空间，往往盘踞在城乡接合部的交通要道，与城市的方方面面息息相关，其内部有着复杂多样的行动者与运作机制。纵瞰全国，大大小小、类型不一的马路劳工市场在各省市星罗棋布。过往研究不在少数，比如山东（刘敏、岳晓林，2020）、东北（田毅鹏、张霁雪、陶宇，2010）和深圳（黄斌欢，2021）。但与本文宗旨不合，因此不加穷举，只是少量提及。吸引笔者选择此处作为调查对象的原因在于，北京市A区日结工市场在政府的干预下从开放马路市场变为了封闭建制市场。这是过往中国研究马路劳工市场的资料所没有观察到的新现象。而且，政府干预形成的建制市场也是过往研究所常常建议的合理形态。那么，这种形态新颖且常被研究者建言献策的日结工市场究竟与其他马路劳工市场有何不同，政策建议的落地又是否会有落差？这构成了最初的疑惑与动力。但这并不是重点。本文不试图标新立异，而是总结与反思“试图标新立异”这一过程中的种种问题。简单来说，本文的目标是成为一个靶子，供读者和笔者自己检讨。因此，首先要做的是立靶子，即整理、组织和描述田野资料。而在日结工市场的诸多现象中，寻工磋商无疑是其中的核心现象。本文将寻工磋商界定为：通过劳动内容与劳动待遇的磋商以确定工作意向和人员选择的匹配活动。它也是该地区马路劳工市场演变的开端。因此，本文将以寻工磋商为主线来组织田野资料，描述经验事实。

二、田野调查资料的组织与描述：以寻工磋商为主线

寻工磋商是日结工市场的核心，是建立在劳动内容和劳动待遇

基础上确定工作意向和人员选择的匹配活动。在如此活动中，日结工市场的市场结构和市场规模构成了基本场域。日结工市场结构指角色结构、工作机会分布状况、市场空间形态和时间模式。日结工市场的规模则指劳动待遇、人群流量与特征等。只有先摸清楚基本场域的框架，才能够理解行动者在场域中的寻工磋商活动，以求能够把握寻工磋商活动中各种微妙与复杂的现象。因此，首先论述的是对市场结构与市场规模轮廓的尝试性勾勒。

（一）日结工市场结构

1. 日结工市场的角色结构

通过分析所收集的田野资料，整个日结工市场内的角色可以被分为三方：日结工、招工者和市场管理员。不同角色有着不同目的与活动，并在市场中占据相互联系的不同位置，从而显现出角色的整合结构。此外，不同角色的内部可能会形成独立的组织结构。

（1）日结工与日结工群体

日结工是该市场内最为常见与多见的角色。简单来说，日结工可以被界定为：以天为基本劳务测量单位与工资结算周期的劳务工作者。他们的目的是赚取劳务报酬。对他们来说，劳务报酬内综合了工资、交通、工作餐、辛劳程度、持续性工作预期和心理愉悦等诸多因素，并非仅仅指一天的工资多少。但这些不同因素大多能通过货币相互比较与综合，从而形成日结工预期工资的可接受范围。

日结工从工作类型可以被分为大工与小工两种角色。大工指具有一定专业技术的工人，小工则指不具备专业技术而出卖劳力的工人。市场内的日结工们对自己的工作类型也有着清晰的认知，自主

将自己划分为大工或小工并寻找相对应的工作。但他们之间的界限并不严格分明。当大工没有合适的工作时，也会揽小工的活。

日结工的活动可被分为三个环节：寻工磋商、实际劳动和工资结算。不同活动发生的场所并不一致。日结工市场内一般只发生寻工磋商活动。日结工们在市场内不断收集信息，与招工者磋商，同他人竞争从而获得一次工作的可能。实际劳动与工资结算则发生在日结工市场之外，但它们会成为日结工的经验参照与未来预期，从而影响市场内的寻工磋商活动。例如，市场内东北招工者的繁重劳动常常使得日结工们望而却步。又或者，现金与微信这两种工资结算方式看似无差别，实际则影响工资差额发放的可能，曾有日结工抱怨，招工者用微信转账克扣工资而自己也无从说理。关于寻工磋商活动，本文将在后续展开描述，此处不多赘叙。

至于日结工角色内部是否存在分化形成的组织结构，从目前的资料来看，难下定论。笔者依稀能够通过观察与访谈中的少量痕迹推断出，日结工市场中的小部分人通过同招工者的劳务磋商来推荐日结工，并从中收取一定介绍费。他们也并不需要进行实际劳动。这意味着，日结工的三类活动都可能分化为不同的角色并有着相应的组织结构。但这些人难以接近。所以，日结工角色内部是否形成团体并具有分化的组织结构？无从论述。

但是，即使可能存在小型的日结工组织，市场内的日结工们在总体上还是比较离散的。他们的内部很少存在稳定且高密度的交往关系，更无须说组织化和制度化的互动关系。他们多以分散的个体形式与其他日结工、市场管理员和招工者打交道。日结工于日结工市场的聚集只是机械意味的空间聚集，而不是有机的组织聚集。日结工中，尤其是男性日结工，就连三三两两的小团体都十分少见。

女性日结工则经常能被观察到抱团现象。此外，为数不多能看见日结工以群体聚集形式出现的时候，往往是当日结工同市场管理组织或招工者发生冲突。但种种冲突中，日结工们虽然大量聚集且相互共情，但组织和制度化程度依旧不高。而且仍有大量日结工旁观冲突，告诫“不要管闲事”。从此判断，日结工们主要以个体形式面对招工者与市场管理员。有趣的是，市场管理员曾不无得意地说过，日结工的刺头已经被他们赶走了。

（2）市场管理员与市场管理组织

市场管理员是日结工市场中的另一重要角色。他们隶属于市场管理组织，主要负责维持日结工市场内部的秩序，以及引导A区日结工市场周围的闲散日结工聚集到A区日结工市场中。他们身着统一的制服，被市场管理组织分为了四支队伍，并安排了相应的队长。每天清晨五点多，日结工市场还处于苏醒阶段时，市场管理员们会被面包车从员工宿舍拉到市场中列队讲话，然后被安排到不同的位置上维持秩序或引导日结工。

从市场管理员被安排的位置来看，整体分布情况呈现内聚外散的状态。内聚指，相当数量的市场管理员集中于日结工市场内部维持秩序。外散指，其他的市场管理员则被零散安排在A区建材市场、地铁站和南关等周围的交通要道上，负责驱散闲散日结工的聚集和引导他们向封闭的日结工市场内聚集。驱散和引导，是A区B街道政府对该市场管理组织的要求。从实际运作来看，市场管理员承担了一定的城市管理职能。因为市场管理员不单单疏散和引导日结工，而且对街道上的地摊商贩也加以管控。曾有街道政府人员在看见街道上的地摊商贩后，通电市场管理组织派人去处理，市场管理员在劝说无效后组成人墙将摊贩团团围住。

从历史来看，A区日结工市场的市场管理组织和市场管理员的历史也比较早。早在马路劳工市场的阶段，由于城市街道同马路劳工市场的种种冲突，A区B街道就已经委托该组织负责该劳工市场的秩序维持。直到2020年的五月份，B街道牵头将日结工市场搬迁。这次，街道主动提供了一块带有封闭围墙的空地，并委托原先的市场管理组织以大巴车将日结工们分批次地运送到如今的A区日结工市场。目前的日结工市场则主要由该市场管理组织负责管理与运作，有着明确的实体边界与功能分区，诸如封闭围墙、区隔围栏、进出大门、办公室、停车场、磋商区、转包出去的商店与早餐店，内部还装有场地喇叭、监控摄像头等设备。

（3）招工者与劳务雇佣组织

招工者往往是日结工市场内日结工所“围堵”寻工的角色。他们的目的是寻找工作能力与工作报价合适的日结工。他们的活动也可以被分为三个环节：寻找工人、劳动控制和支付工资。在日结工市场中，我们也只能观察到招工者的寻找工人活动。惯有的工作搜寻模型往往认为劳动者要在市场内不断搜集工作信息并为之付出成本（乐君杰，2011）。但实际上，我们不难发现，日结工市场中的招工者同样要不断搜集工人报价并与自己的心理价位相比较。双方是一个互相匹配的过程。工资的分布并不是外生既定的，而是双方磋商的结果。当匹配不成功时，双方都有损失。日结工市场内既有找不到工作来维持生活的日结工，也有找不到工人傻站着的招工者。

整个日结工市场内的招工者可分为两类型，临时招工和固定招工。临时招工指的是，这些招工者并不会以确定的周期在市场内招工。这些工作主要来自A区周围或北京市内小微型家装需求，以及

建筑工地或绿化队的清理与绿化工作。固定招工则指的是，持续性地稳定在市场内招工。这些工作大多来自小型建筑队和快递公司等，比较出名的是市场内的东北建筑队招收瓦工[1]。总的来看，除少部分小型建筑队和快递公司外，招工者很少是高度组织化和制度化的群体，或群体中的成员，他们的用工需求比较随机。他们内部可能有着独立的组织结构，但由于调查地点的限制，未能深入观察。

2. 日结工市场的工作分布

北京市是一座巨型城市，内部复杂多样，充斥着种种不同类别和阶层的人群与不同类别和层级的需求。在现代化、城市化的北京，劳动分工的进程十分迅速。建筑和家装需要相应的专业劳动技术。市区内同样有着许多小微型的建筑安装需求，例如北京市周围村庄的修缮、重建工作，小型店铺或家庭的家装需求等。由于成立一家正规建筑公司有着诸多制度要求与经济压力，例如注册成本、运作成本、税收成本、各种保险、工会会费等等。与此相匹配，雇佣一家正规建筑公司进行施工的成本也并不低。对于小微型工作需求而言，他们不愿支付正规建筑队或家装队的高雇佣成本，雇佣日结工的性价比相对较高。而且，北京市内有着大量的建筑工地与绿化工程。在建筑工地上清理建筑垃圾、搬运建筑材料，或者绿化除草等工作需求量也十分大。对于建筑队来说，正式雇佣劳动力进行低技术工作的成本很高。以上种种构成了日结工市场内部工作机会的来源，并形成了相应的分布结构。

---

①日结工市场内有着不少操着东北口音的招工者长期来这招瓦工，统称东北老板。但他们常被日结工们诟病用人太狠导致第二天无法工作。

因此，整个日结工市场中分布的工作多与建筑行业相关，可被分为技术工作与低技术工作。市场内的技术工作主要可分为瓦工、木工、水电工和焊工，细分则有砌墙、抹灰、支模、贴瓷砖、吊顶、电焊、家装等。技术工作大多要求携带专业的劳动工具，因为此处的工作很少能够准备专业工具。是否拥有专业工具是工人们展现工作能力以争取工作的重要手段。低技术工作则一般无需具备专业技能，并不要求携带专业劳动工具。它们主要是些出卖劳动时间和工人体力的工作，诸如送砖送水、打墙、拆除、除草、清理垃圾、搬运等等。总的来看，该市场内的工作大多是低技术工作。从访谈资料粗略分析，原因有二：其一，建筑工地一般雇有自身的技术工人，来市场只是招勤杂工；其二，A区日结工市场的辐射范围有限，城市中大量非正式的技术用工需求并不是只有此处可以满足。

3. 日结工市场的空间形态

整个寻工磋商的空间特征可以分为两个时期，即开放的马路劳工市场时期和封闭的A区日结工市场时期。在调查开始前，开放的马路劳工市场就已经变化为当前的封闭空间。因此对它收集的资料仅限于少量的口述，甚至可以说是空白，因此难以描述。而且，资料的空白不仅仅由于变化已经发生，更是由于调查的缺陷，这点待后续详谈。

马路劳工市场的空间特征可以被总结为开放、混杂和扩张。从口述资料来看，最开始形成于A区建材市场周围的马路劳工市场并没有明确的市场边界，只是由许多的工人聚集在马路边或大空地。只要有一定数量的工人聚集，这地方就构成了一个劳工市场。边界

不明的情况下，马路劳工市场形成了大小聚集点混杂的情况。大量工人聚集在南关、地铁站或建材市场附近，少部分工人聚集在建材市场周围的各大道路旁。宽泛地说，A区建材市场的周围都可以是马路劳工市场。

这种形态也就意味着，原先的马路劳工市场并没有明确的物理边界同城市环境相区分，只有当时的市场管理员可在一定程度上被视为标定劳工市场的符号边界。由于活跃时间与城市时间的不一致和对空间边界的不确定，马路劳工市场频频与城市中其他系统发生冲突。它不断在城市公共空间中扩张。给城市公共空间带来了很大的压力。当地政府（街道）对此十分头疼。工人的生存、生活和情感，政府人员的引导、管理和整治，以及居民的抱怨、愤懑和投诉，都在这之中被层层折叠成马路劳工市场的日常。随后在多年的各种力量碰撞中，马路劳工市场才在政府牵头下变成了现今的A区日结工市场。

现今的A区日结工市场的空间形态与马路劳工市场时期相比，发生了很大的变化。其中最为显著的变化，就是市场边界的明确和内部功能区的划分。现今的A区日结工市场远离原先以建材市场为核心的闹市区，搬迁到人流稀少的大道旁。它以涂有蓝漆的低矮灰墙加以封闭，并由一道市场管理员把守的大铁门控制进出。市场内部划分有不同的功能区，诸如交易区、休息区、办公室、停车场等等（图一）。市场管理员要求日结工们在不同功能区中各得其所，但往往事与愿违。总需要市场管理员们软硬兼施来维持各功能区的基本分隔。日结工们总是热切地想要穿过规定的交易区去往停车区，围住驶进停车场的各种车辆，试图寻求到一份好的工作。封闭的空间让他们觉得，潜在的招工者会由于围

墙的遮蔽而忽略这里辛勤的劳动者们。他们想要回到从前，但是市场管理员不会允许。

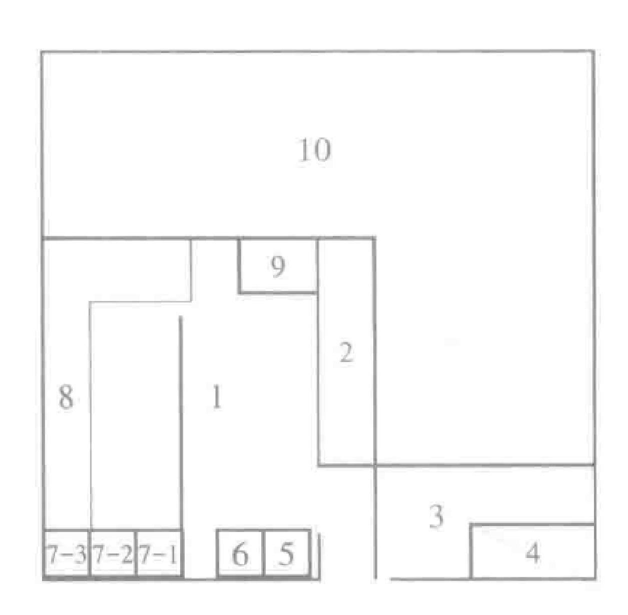


图1 A区日结工市场

4. 日结工市场的时间模式

空间区隔往往是系统边界的直观显现，使得区别于环境的系统得以“现形”。在卢曼那，系统的时间制度是系统化约复杂性的重要机制（郑作彧，2022）。关于时间的社会学研究，主要处理时间的社会性与社会的时间性两个问题，但常陷入矛盾。在索罗金和默顿关于时间的研究中，社会活动与社会时间往往紧密相连，社会时间为集体活动提供了同步与整合的工具，且会随着社会结构变化（Sorokin & Merton，1937）。这一观点主要强调了作为社会制度的时间制度，即时间的社会性。随后的杰鲁巴维则强调社会的时间性，认为不同角色和组织都需要创设错开或并行不悖的时间制度才能将人纳入其中，使得组织与角色得以成立与运作（Zerubavel，1979）。此处并不意图梳理和解决关于社会时间问题的研究脉络和疑难，而

是想强调要调查清楚研究对象在时间维度上的特性。但许多研究往往将研究对象的时间维度视作其在物理时间上的展开，而非深究其展开中所显示特性的社会学意义。

日结工市场中，在物理时间轴上展开的寻工磋商活动在总体上显现了时间上的周期性特征。简单来说，日结工市场具有时间性特征，并呈现特定的时间模式。A区日结工市场中，少部分日结工大约在清晨四点多便会零散地来到场地内。清晨五点到五点一刻左右，市场管理员们则会开始列队训话。随后的五点一刻后，日结工们开始陆续进场。日结工市场开始复苏，人气逐渐旺盛起来。招工者们在这个时候也逐渐进场招工。五点半到六点半，是人气最为旺盛的时候。这个时间段内，进出的车流与人流都十分密集庞大。到处都能听见日结工们的招呼、闲谈与抱怨，招工者们的呼喊和各种寻工磋商的言谈。它们或激烈，或平和，或悄无声息地招一下手人就跟着干活去了。到了六点半，市场内的人气开始盛极而衰。许多招工者也逐渐结束招工。因为在很多劳务活动中，一天的工作往往从六点半开始计时。到了六点半后，市场内的工作便开始减少。此后进场的招工者多半零散，招人也不多。但他们每次进场，都会引发日结工们的围堵，就像浴室的地漏漩涡式地吸引着所有。拥闹的招工漩涡与平淡的散居休息在市场内互相交叉。间歇性的聚集高潮反复上演，又反复落幕。这将一直持续到上午的八点至九点。八九点后，日结工市场则陷入更长的休息，因为寻工磋商活动已经进入了尾声。下午六七点钟，部分被车辆载去工作的日结工们又会被拉回到市场，作为一天工作的结束。因为日结工们大多住在A区日结工市场周围廉价的自建房中。

在描述日结工市场内一天的周期性变化后，我们不难发现日结工市场的时间特性与其所面临的社会环境有着密切关系。往大了

说，日结工市场的发轫与壮大同北京市过去多年的快速城市化和大范围城市建设相关[①]。往小了说，日结工市场中人群聚集的开始与结束同市场所分布工作的时间要求有着密切联系。此外，开放的马路劳工市场之所以演变为如今的封闭市场，也很大程度是因为同城市公共空间的时间安排发生了冲突。交通要道的时间本是用来供车辆流动，但日结工也占据交通要道的时间来等工找工。从上文对日结工市场内一天的描述[②]也可知，城市居民的休息时间常与日结工们的找工时间相冲突。日结工和居民对各自时间的规划原本构成了无形的时间边界以区分不同活动，但由于处于同一空间且不能和谐共享而致发生不少矛盾，最终由政府划出封闭空间使得二者得以通过物理边界相互分隔。

（二）日结工市场规模

结构是市场的骨架，规模则是包裹骨架的血肉，二者共同形成了市场的种种特征。但是，对于日结工市场内的性别分布、年龄分布、教育分布、地域分布、工作机会总量、人流进出和劳动待遇等量的特征，未进行数据收集和统计推理。此处只能够基于现有访谈材料和市场观察进行推测。必须诚实地承认，这是严重失误。但是，也需将现有资料诚实展出。

就性别而言，男女的分布状况比较明了，女性日结工虽有一定数量，但相较于男工，还是少数。而且，几乎所有女工都是小工，并不具备专业技术。市场中的女工竞争力也不强，常有招工

①详见北京市统计局的《北京统计年鉴》关于常住人口、房地产开发投资、建筑安装投资、建筑业产值等历年快速增长数据。此处因篇幅所限，不加展示。

②时间的周期性变化上，马路劳工市场与A区日结工市场是一致的。因为其所面临的社会环境基本稳定。

者排斥女工。适合她们的工作多是除草、清理垃圾。因此，女工内部的竞争也很激烈。就年龄分布、教育分布和地域分布而言，观察和访谈的日结工的年龄普遍在45岁以上，少有1980年后出生的新生代农民工。他们大多是初代农民工。一个可能的佐证是，青年人在日结工市场比较受招工者欢迎，具有较强工作竞争优势。但招工者又往往难以招到青年日结工。而且大部分日结工的教育程度并不高，小学或初中学历居多。高中学历算是比较高的学历，而且足以令日结工自豪[①]。就流出地域而言，他们大多来自华北和东北地区。其余地区少量分布。至于工作机会总量与人流进出量，并不能给出一个量上的判断。

至于劳动待遇，则有较多的访谈资料供以交叉检验。我们在上文提及，整个日结工市场中的工作主要分为技术工作与低技术工作，日结工们也可被分为作为技术工的大工与低技术工的小工。小工与大工的收入差别较大。大工的工资最低也是350~400元/天，具体工资还得视具体工作的复杂程度和劳累程度而定。而且，如果招工者报价低于350元，大工一般会拒绝。此外，市场内有着不少瓦工。他们的工资主要是计件工资制，一块砖0.15元。同瓦工的交流来看，一个瓦工一天可以砌3000至5000块砖。也就是说，一个瓦工一天的工资收入是在450~750元内浮动。但相应地，瓦工是一项比较劳累的工作。对瓦工来说，砌4000块砖对身体来说已经是很大的负担。小工的工资则没有大工那么高。小工的工资大致在200~300元/天之间。而且，随着时间的变化，小工的工资也会随着供需关系发生变化。等过了招工的高峰期，雇主不多，工作机会不

①一位场内工人非常骄傲地说自己是高中生，与其他人都不一样。其他人证实了此点，但也显露出不喜。

多，但是小工数量依然可观的时候，小工的工资则会下降，大致在200~250元/天之间。这与大工完全不同，因为大工的议价并不会随着招工高峰期的过去而降低。据访谈，大工一个月大多能工作二十多天，到手八千多。小工一个月也能有五六千元收入。而市场管理员的收入相较而言，则略微较低。市场管理组织长期在日结工市场内招募市场管理员，要求年龄在18至55周岁，一个月薪资为3500~5000元，且包吃住。市场管理员中也有些原先的日结工，因此与场中的日结工十分熟悉。他们见面大多会打招呼与戏谑一番。发生了冲突，这些市场管理员也是平息双方情绪的润滑剂。由此看来，日结工与市场管理员之间的隔绝并不大。在抛去制服和角色的符号意义后，类似的个人特征使得双方也能够互相理解难处。正如一名日结工与所说："他们也有他们的难处，都是来讨生活的。"

（三）日结工市场中的寻工磋商现象

大致勾勒市场结构和市场规模的轮廓后，我们则可以开始描述日结工市场中的寻工磋商现象。此外，上文的角色结构部分只分别论述了日结工、招工者和市场管理员三方角色，却并没有详细论述三者的互动以及稳定的整合结构。但倘若单单论述角色之间的互动模式，难免受到机械论的批评。这种批评总是如此口吻：将复杂的人类现象视作无意识的机械模型，人成了名为角色的提线木偶。此类批评有着一定道理，但也失之偏颇。诚然，在实际观察中，观察者根本看不见任何单一呆板的角色，反而是形形色色的行动者充斥眼帘。五花八门的行动现象让观察者眼花缭乱，随时可能陷入失焦的陷阱。但是，也正是从繁杂多样却暗含逻辑的各种现象中凝练出的角色能让我们聚焦逻辑从而一探全貌。换句话说，角色是观察者

对行动者复杂性的凝练，有助于我们把握特定场域中活动的结构限制。因此，我们一方面既可以通过角色理解行动者，另一方面我们又可以通过行动者的能动活动揭示角色之间的互动模式、整合结构与结构变化。上文我们已经通过勾勒各方角色来刻画理解行动者的参照系，接下来将通过描述寻工磋商现象中各方行动者的能动活动来揭示角色的互动模式、整合结构与结构变化。这样的话，既避免他人的机械论批评，也避免被各式行动者迷花了眼。

1. 寻工磋商—招工漩涡

在角色要求、时空制度和工作分布等市场结构限制下，有着不同惯习、个人特征与过往经历的行动者们在日结工市场中有限制地能动活动并形成繁多现象。正如马克思所说："人们自己创造自己的历史，但是他们并不是随心所欲地创造，并不是在他们自己选定的条件下创造，而是在直接碰到的、既定的、从过去承继下来的条件下创造。"（马克思，2009：470—471）

招工漩涡是对某类寻工磋商现象的形态描述，指日结工们涌向招工者所形成的层层包围聚集形态仿若浴室地漏吸纳流水的漩涡（图二）。这种形态上较为特殊的招工漩涡现象比较集中地展示了日结工市场内的主要活动与矛盾，涉及多方行动者在场域中的角力。因此，通过描述招工漩涡可以折射出：随着工作机会在时间维度上的分布变化，封闭空间内不同行动者的互动呈现何种形态，这种互动形态又如何披露与冲击角色之间的稳定结构。简单来说，招工漩涡是理解日结工市场繁杂现象的枢纽。

接下来的笔墨，将集中用于描述一次招工漩涡现象。但是，每次对招工漩涡现象的观察往往只能揭露现象中的某一特征而非全

貌。将整个现象中所有异质性要素显露出来的典型案例在经验上可能并不存在，又或者由于研究者调查水平与调查期限等偶然因素限制而未能观测到。所以，在此并不能够通过单一案例的描述来全面展示招工漩涡现象。但出于全面展示招工漩涡现象的考量，又对招工漩涡现象的描述糅合了多个经验案例。笔者将从多个经验案例中抽取不同的异质性要素共同塑造出一个在经验上尚未被观测到的典型现象。这一做法参考了韦伯的历史案例（historical case）概念（韦伯，2010：25—26）。

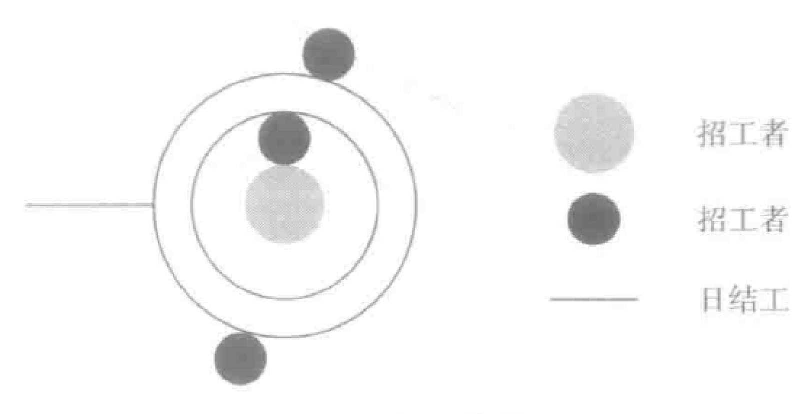


图2 招工漩涡

在一次典型的招工漩涡现象中，随着时间的推移，日结工市场内的工作机会总体减少且不再集中出现。这时，招工者A开车来到日结工市场。他下车后，清晰地当众声明招工意图。少数日结工见状上前询问工作的具体内容与待遇，并展示自我的资历、工具和技术水平。双方开始磋商讨论具体事宜，形成了一个有序的小型互动系统。场内的其余日结工看见A同少部分日结工聚集起来讨论工作事宜，自发地被吸引到原来的小型互动系统中。他们同样希望能够得知工作信息，争取一个好的工作机会赚钱维持家用。然后，日结工们就像水流般涌入以招工者为核心的互动系统中。A被日结工们层层包围了起来。小型互动系统扩张的同时变得十分拥挤杂乱，人

声鼎沸。不断加入的新日结工们反复地问询A工作内容和工作待遇等事宜。为了能在拥挤的人群中吸引招工者注意力，被卷入漩涡的新日结工们问询的声音不会太小。更有甚者，直接不问具体工作与待遇，拍拍A就扭头指着停车场说"就我了，我跟你去"。这种情况下，A要么不断重复简单明了的工作内容和工作待遇，要么一概不理，仅同最初的日结工们交涉更为详细深入的工作内容与工作待遇问题。如果招工者从未见过如此混乱嘈杂的阵仗，少部分被人群纠缠得紧了，还会推开日结工所组成的"人墙"，连忙摆手说"不招了，不招了"，然后就离开。这足见其中的拥挤与嘈杂。

但是，拥挤嘈杂归拥挤嘈杂，正经事不能耽误。A与日结工们的劳务磋商得进行下去。A要招工人做事，日结工要找工作赚钱。在招工漩涡中的劳务磋商中，双方既要相互交换信息，又要不断比较自己的心理价位和对方出价。通常来说，信息交换是对等的。不仅仅日结工要获得工作信息来衡量劳力投入和工资收入，A也要说明白工作信息以免活做不好。对于技术工种而言，这尤其重要。日结工会不厌其烦地问具体工作细节，比如怎么走线、焊接点有哪些。工作细节会成为技术工同A展现能力来讨价还价的重要倚仗。工作细节不明晰的技术工作，日结工不会妄下论断揽工，但也不愿放弃，会跟着A去实地看看情况再做决断。但有时信息交换也并不对等。如果A有意隐瞒具体工作信息试图以较低工资招募工人，工人也不傻。一旦看对面支支吾吾或在追问中露出马脚，发现工作内容与工作待遇并不相符。他们也会你一言我一语地要求加价。日结工们都要求加价的群体压力，对A也是不小的心理负担。还有一些日结工并不挤进招工漩涡的中心，反而总是在外围评论这工作的待遇与内容，有时也给招工带来不小压力。如果A坚持不加价，漩涡

则会很快散去。A只能在市场内游荡却无人问津。如果A在一段时间的无人问津后表示加价，新一轮漩涡与讨价还价则又会开始。

如果工作待遇与工作内容比较合适，这份工作便是一个香饽饽。日结工们会与A称兄道弟，还会展现自己的专业技术过硬、工作经历丰富而且还吃苦耐劳。假若A要招募的工人比较多，少部分日结工还会很有领袖气质地大手一挥说："甭麻烦了，我帮你找剩下的人。"这少部分有领袖气质的日结工既能找人，也能帮A控制住杂乱的热闹场面。然后他会拉来自己熟悉的朋友一起去做工，其余想做工但不熟悉的日结工们则开始和这名日结工攀交情。如果他朋友人数不够，他就会快速地在招工漩涡中补齐空余，然后跟着A上车。但到这还不算结束，少部分不甘心的日结工即使看到A已经找好了工人，也不愿意放弃。他们会继续跟到停车场，试图揽到这份工。更有甚者，直接坐进车内要求再招他一个。这时又会发生漩涡式的聚集，但不同的是其余日结工是过来看热闹。

当然，以上过程中，市场管理员也不会闲着。总会有一个市场管理员开着个大喇叭挤到A身边，帮着组织秩序和传达工作信息。还有些市场管理员在漩涡外围维持秩序，防止生乱。假若这招工漩涡发生在停车场，市场管理员还会驱赶A与日结工们至交易区。不少日结工和A会因此不满。见惯了漩涡乱象的市场管理员因此会主动将招工者直接拉入交易区。招工漩涡的日结工们与市场管理员有时会由于各种偶然因素发生冲突，比如管制言语带来的争吵、一次偶然的肢体接触等等。当双方发生冲突时，日结工总觉得市场管理员妨碍了自己找工作，组织秩序的言语粗暴且各种规矩不合理。市场管理员则觉得维持秩序是自己职责，日结工们太乱太吵。话虽如此，在一般情况下，市场管理员依然能够组织和维持招工漩涡中日结工们

的秩序以免太过混乱。他们的手段也并非总是强硬的话语与驱赶，也讲究个先礼后兵，软硬兼施，例如称兄道弟或情感诉苦。当漩涡中的日结工们关于工作的竞争过于激烈时，市场管理员也能起到缓和的作用。有时，市场管理员会帮A在人群中确定工作人选以稳定秩序，然后迅速地让他们乘车离开。除此之外，市场管理员一直在思考如何更有效地维持秩序，例如在交易区内设置物理区隔将招工者与日结工分离。但这一做法的效果并不是很好，跨越区隔总是无可避免。显然，市场管理员并没有将日结工们的想法考虑在内。

在上述典型现象中，我们可以尝试着从复杂性中凝练出如此叙述[①]——随着工作机会在时间维度上的分布变化，行动者之间的寻工磋商呈现漩涡状聚集形态的可能性加大。在如此漩涡中不难发现，日结工的内部高度离散，其互动多是相互竞争，组织化与制度化程度不高。日结工中可能存在小团体但为数不多。日结工们大量聚集，引发市场管理员介入维持秩序。市场管理员与日结工、招工者之间呈现二级的扁平管理结构，但管理效率低下，管理威权模糊。市场管理员与日结工们不时由于偶发因素发生冲突，多次冲击管理结构。但该管理结构却依旧显示出韧性。招工者与日结工则是形式上平等的磋商互动，在磋商中搜寻和交换工作信息（如工作内容、工作待遇和工作时长）与人员信息（过往经历、技术高低、专业设备、人品性格）以求互相匹配。磋商过程既有严肃的讨论，也有游戏的轻松。只不过，寻工磋商的信息交换效率比较低，秩序比较乱，常常冲击市场管理员的组织与管制。

①描述与叙述有着很大差别。描述是对事件复杂异质性与异质性内部逻辑的尽可能描摹。叙述则往往加入了情感、理论、立场约束复杂的异质性并解读其内部逻辑。在实际调查中，我们只能观察和收集叙述。描述是建立在不同叙述基础上的研究工作。

2. 寻工磋商—零散招工

如果说招工漩涡在形态上是高度集中的寻工磋商，那么零散招工则是形态上较为离散的寻工磋商。依旧采用上文典型现象的塑造办法，从多个经验案例的糅合来描述零散招工的典型现象，以期描述清楚另外一种不同形态的寻工磋商现象，来加强对日结工市场内能动活动的把握。

在一次典型的零散招工现象中，招工者B会挑一个好时段来到日结工市场。他并不会大肆宣扬自己要招工。他已经是个招工老手了，经历过多次招工漩涡，也了解市场内的行情。他不想被那么多人围着，效率低且吵闹。他就像一个经验丰富的猎手在日结工市场慢悠悠地负手漫步，就好像是来看热闹。他甚至看着像是来找工的日结工。这也很有可能，因为他确实可能是受老板信任和委托来市场招工的日结工。他不怎么着急，因为他可能在等待已经通过电话联系好的日结工、又或者他只想找市场内的熟面孔去工作。反正他不着急，可能还会和熟悉的日结工开始谈起天来。市场管理员也不会注意他和管他。因为他既不引发漩涡，也在规定的交易空间内活动。又或者他突然眼神一亮，开始问一个刚来的熟人找不找活，又或者找熟人来问他有没有活。但也有可能当其他工人来问工时，他摆摆手说自己不是来招工的。他和日结工们谈工作的时候很是老道，知道什么工作该给什么待遇，不慌不忙地介绍工作信息与工作待遇，甚至给他们点上一根烟，叫上一句亲切的老弟或老哥。如果看到年轻点的工人，他也会上前问找活不。他有一套自己的选人方式，比如能干、年轻、信任、熟悉等等。不快也不慢，他就要召集齐人走了。

很显然，零散招工的形态及所透露出来的种种痕迹都显露出其

不同于招工漩涡现象。我们依旧可以从中凝练出一条叙述——零散招工形态下的互动仅涉及招工者与日结工，市场管理员并不介入。因为零散招工发生冲突的可能性不大，所以不太可能扰乱秩序与动摇管理结构。或者说，零散招工默认了市场管理员的存在，自觉遵守管理秩序以避免不必要的麻烦。此外，在他们的互动中，过往业务带来的熟人关系痕迹重。不同于招工漩涡，其招工效率高，竞争强度低，秩序较稳定。而且，招工者隶属某一劳务雇佣组织的可能性高，其内部有着制度化的组织结构与稳定的招工需求。

综上，不难发现，招工漩涡与零散招工是日结工市场的两大缩影。二者既浓缩着角色结构、工作分布、时空制度等结构限制，也透射着行动者的性格、惯习、经历与能动策略。我们可以看见，复杂异质的现实生活与人类意图和行动之间存在种种限制与超越。一方面，现实生活如此复杂异质，不同的潮流与力量驱动着和限制着人们迈入不同方向。另一方面，每个人都试图在生活激流中利用或对抗潮流的裹挟，以求保持自己的平衡与稳定，从而实现自己的意图。其中的张力将前文略显僵硬的结构描述生动了起来，也留下了进一步探究的可能。

三、坦诚的反思：课堂与田野的距离

（一）调查区域的选取

复盘日结工市场的调查，深感选定一个调查区域并不简单。它是一个受到多方面因素共同作用的复杂决策。这一方面是由于完整调查的客观要求，比如时间因素，调查区域的大小与调查时间的长短往往呈现正比例关系。调查区域越是广阔，其中的异质与矛盾则

更为丰富，耗费的时间自然更多。另一方面，能力程度不同的调查者所能处理的信息量也不同。调查区域与自身能力不相适配，调查最后也难以达到满意的结果。此次的田野调查就是如此。就像儿童走进成年人的复杂世界，一切都成了迷雾与谜团。不单单各种问题令人困惑，各种问题的解释也更是如此。而笔者在此中兜兜转转，注意力不断聚焦、失焦。

而且，由于调查者在社会结构中本身也占据一定的位置，承担一定的角色。在这个位置与角色之中，调查者也面临不同方向的压力。比如，开始和结束调查的决定权有时并不在调查者手中，而受上级经费规模、时间限制、政治考量甚至个人好恶等外在因素影响。此外，研究者角色这一身份也期待研究者能够通过调查提出学术问题、推进学术发展。而且，调查者仅拥有相对程度的完整信息，往往要通过新闻媒体、相关研究、相关引导人等渠道来获得预调查资料，并以此作为选定调查区域的依据。但预调查资料与真实情况既可能相差不大，也可能天差地别。但在多种压力逼迫下，选定调查地点后，实地中即使发现决策失误，也只能硬着头皮进行下去。

最后，对调查区域也有一定要求。调查区域既不能大而无当，也不能过于微小。最重要的是，选取的调查区域必须能够包含足够的异质性要素，能从中透视多种社会事实、社会制度与运作机理，即能够见微知著。

总结来看，选定调查点并不是一个寻找逻辑最优解的过程，而是寻找实际可行解的复杂过程。此次田野调查也不下定论，而是通过调查给出丰富的描述资料，以期同其他不同类型的劳务市场相对比，为其他研究者更为详细与深入的知识论证提供先行。这与费孝通组织江村资料并定位研究位置的做法是一样的（费孝通，1986：6）。

（二）社会学理论与实际田野

调查过程中，有一难题常让人困惑——如何认识和处理社会学理论与实际田野的关系？对这一问题的回答，常有两种论调，一是累赘论，二是拐杖论。累赘论认为，理论会预设和限制田野实际情况，使之成为经典理论的复现，所以最好是脑袋空空下田野。拐杖论则认为，理论是辅助研究者从独特学科视角观察田野实际的拐杖。理论不是关注多了，反而是关注少了（Burawoy，2013）。而本次田野调查在这两个论调上都栽了跟头。

起初是脑袋空空下田野，一门心思想着观察所有看得到的田野现象并如同照相机一般将他们全部复刻下来。但是，什么都看的结果就是什么都看不到。就像刘姥姥走进了大观园，所有东西都新奇得不可理解。所有异质性全部摆在了眼前，却根本无从组织和简化。焦点消失了，整个人被复杂的现实情况吞没。在观察日结工市场的同时，却忘却了自观察的重要性。最终的后果就是收集了一堆支离破碎的观察碎片。这是走入的一个极端。

意识到此点后，又开始埋头理论，尝试着以理论作为事实的透镜去重新开展调查。但又走入另一个极端。一头扎向了劳动力市场分割（赖德胜，1996）、就业分流（万向东，2008）、城市融入（王春光，2001；陈映芳，2005）、分层流动（李强，2003）和阶级抗争（潘毅、卢晖临、张慧鹏，2010）等多个理论主题，然后看看是不是能碰运气找到解释调查资料的理论。这样做的结果也很明显，眼里只会剩下种种符号。一大佐证就是，上文的资料整理中，招工者部分体量的简略不是出于详略得当的行文考虑，而是调查根本没有意识到要对他们同等关注。在此，招工者成了符号化的阶级，潜

意识的背景板。根据结果挑选样本，这显然不公道（马丁，2020：45）。那么，走入两个极端的症结究竟在哪？

第二个极端的症结在于研究中倒转了实存与认识的关系。笔者并不是将实存作为认识的开端，而是将认识作为实存的开端。在田野中找农民工、阶层和市场分割，却唯独忘记了确立日结工市场本身。在解释社会事实因果链、因果结构之前，首先需要做的是确立真实的社会现象（establish phenomenon）（Merton，1987）。笔者急着去解释，却忽略了描述的基础性与重要性。更有甚者，将描述性理论归结为低层次研究工作，给理论划分出贵族式的等级高低。如此的后果只能是用虚无的学科术语来构筑漂浮的铁笼然后在其中沾沾自喜。

那么在田野中所观察和描述的真实社会现象又是什么呢？实际就是社会现实具有稳定特征的恒常与生活逻辑（杨善华、孙飞宇，2015；折晓叶，2018），常以制度、系统、结构等为成熟形式。它们是人互相作为观察者以及互相知晓对方是观察者下逐渐形成的稳定互动模式。博弈论均衡可以作为理解它的简化模型。古代政治学常说的“君舟民水”，也就是行动者在行动前观察到其他观察者在观察他后所总结的均衡理论，正所谓“决定者和受波及者的区别将我们带到二阶观察的层次”（卢曼，2020：312）。在这个意义上，描述性理论则更具有元理论性质。描述性理论的模糊也反映着元理论范式的不确定性。

但如何进行描述性观察并整理出描述性理论呢？这是第一个极端的症结，少有关于描述性理论的方法论反思与研究范例来帮助聚焦田野中该看的和忽略不该看的。这也是本文中删去大量的理论与经验文献、重新组织田野资料与描述经验事实的初衷。而且，在这个工作中，质性与量化的区别毫无意义，都是描述的重要维度。关于量特征资料的缺乏，是描述工作中难以被饶恕的失误。

四、结语：迈向可能会更好的明天

本文组织和描述经验材料的方式显然是从系统和整体出发的分析路径。而且只能够描述当下的构成，却无法描述历史的形成。日结工市场内不同结构的相互关系并未能清晰地说明，而只能借助劳务磋商活动的不同形态描述来暗示暧昧的可能。这种可能在于：即使不同行动者受到种种结构约束，但他们的能动活动依然绽开了新可能，以及显露了旧有结构生成的依稀痕迹。这为进一步思考与讨论打开了空间。但这一切都需建立在坚实的描述基础上。只有描述清楚，才能够把握现象的真实从而比较不同真实现象来厘清其中本质的边界，并以此去抽象倒推现实的历史建构逻辑。这种抽象的过程类似一个倒金字塔，从底部尖端不断抽象到顶部。而建构过程则是从顶部的无数可能发展至底部的有限可能。这一可能性展开与约束的过程和逻辑，正是我们所要研究的关键。明天可能会更好，也可能会更坏，但迈向明天的方式可以通过昨天来理解。

参考文献

费孝通，1986，《江村经济》，戴可景译，南京：江苏人民出版社。

国家统计局，2020，《2019年农民工监测调查报告》，http：//www.stats.gov.cn/tjsj/zxfb/202004/t20200430_1742724.html

黄斌欢，2021，《从何而来的“大神”：日结体制与悬浮社会》，《社会发展研究》第4期：39–57+240。

赖德胜，1996，《论劳动市场的制度性分割》，《经济科学》第6期：19–23。

乐君杰，2011，《工作搜寻理论、匹配模型及其政策启示—2010年诺贝尔经济

学奖获得者研究贡献综述》，《浙江社会学科学》第1期：135–140+159。

李强，2003，《影响中国城乡流动人口的推力与拉力因素分析》，《中国社会科学》第1期：125–136+207。

刘敏、岳晓林，2020，《"有市无场"与"有场无市"：马路劳务市场及其治理困境》，《南京农业大学学报（社会科学版）》第5期：38–47。

马克思·韦伯，2010，《新教伦理与资本主义精神（罗克斯伯里第三版）》，苏国勋、覃方明、赵立玮等译，北京：社会科学文献出版社。

尼克拉斯·卢曼，2020，《风险社会学》，孙一洲译，南宁：广西人民出版社。

潘毅、卢晖临、张慧鹏，2010，《阶级的形成：建筑工地上的劳动控制与建筑工人的集体抗争》，《开放时代》第5期：5–26。

田毅鹏、张霁雪、陶宇，2010，《空间生产、资本接续与权力介入的实践逻辑》，《社会科学》第5期：67–73+189。

万向东，2008，《农民工非正式就业的进入条件与效果》，《管理世界》第1期：63–74。

王春光，2001，《新生代农村流动人口的社会认同与城乡融合的关系》，《社会学研究》第3期：63–76。

杨善华、孙飞宇，2015，《"社会底蕴"田野经验与思考》，《社会》第1期：5–26。

约翰·李维·马丁，2020，《领悟方法：社会科学研究中的方法误用与解决之道》，高勇译，重庆：重庆大学出版社。

折晓叶，2018，《"田野"经验中的日常生活逻辑：经验、理论与方法》，《社会》第1期：1–29。

郑作彧，2022，《时间的系统构成——卢曼社会系统理论中的时间概念》，《社会学研究》第2期：69–91+227–228。

Burawoy, Michael 2013, "Ethnographic fallacies: reflections on labour studies in the era of market fundamentalism", *Work, employment and society* 27(3): 526–36.

Merton, Robert K. 1987, "Three fragments from a sociologist's notebooks: establishing the phenomenon, specified ignorance, and strategic research materials." *Annual review of sociology* 13(1): 1–29.

Sorokin, Pitirim A. & Robert K. Merton 1937, "Social Time: A Methodological and Functional Analysis." *American Journal of Sociology* 42(5): 615–629.

Zerubavel, Eviatar 1979, "Time and Public Time: The Temporal Structure of Social Accessibility and Professional Commitments." *Social Forces* 58(1): 38–58.

冯军旗．新化文印产业研究的过程和思考[M/OL]//赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

新化文印产业研究的过程和思考

冯军旗[1]

摘要：本文通过叙事的方式，回顾了从事新化文印产业研究的缘起和过程，并具体考察了参与观察、问卷调查和深度访谈等方法在研究中的具体运用，同时，对于研究中遇到的障碍和挑战以及解决的办法，也进行了具体的阐述。最后，文章还对新化文印产业的演变和发展进行了思考和反思。

关键词：新化文印；参与观察；问卷调查；深度访谈

Abstract: Through narrative, this paper reviews the origin and process of engaging in the research of the Xinhua Wenyin, and specifically examines the specific application of methods such as participant observation, survey and in-depth interview in research. At the same time, the obstacles and challenges encountered in the research and the solutions are also elaborated. Finally, the article also reflects on the evolution and development of Xinhua Wenyin.

Key words: Xinhua Wenyin; Participant Observation; Survey; In-Depth Interview

①作者简介：冯军旗，男，1976年生，河南人，博士，中国政法大学政治与公共管理学院副教授，硕士生导师，联系方式：kant222@126.com。

一

我在2006年做新化文印产业的研究，纯属偶然。当时，我在北京大学社会学系攻读博士学位。我的导师郑也夫教授给我们开了《城市社会学》的选修课，要求每位同学以城市里的一个社会现象为主题，进行研究，并提交一篇5000字左右的课程论文。我当时由于喜欢读书，经常从北京大学图书馆借一些书交给北大内的复印店复印。时间长了，我发现这些开复印店的人基本都是一个地方的人，都说一种极其难懂的地方方言。他们当时也经常给我说，北京开复印店的基本都是湖南新化县的，不少都是洋溪镇的。在这种情况下，我决定就以新化人的文印产业作为课程论文的主题，当时主要是以北京大学校内的复印店为研究对象，以深度访谈和参与观察为主要研究方法。大约进行了一个多月，我就提交了5000字左右的课程论文。但是，以严厉著称的郑老师很快给我当头一棒：调查很不深入，重做。在这种情况下，年轻气盛的我痛下决心，一定要好好地做一做，争取交给导师一个像样的东西。

随着研究的深入，我发现新化文印产业已经有了几十年的历史，经历了多个发展阶段，并呈现出层累式的螺旋上升态势。这个产业当时已经形成了遍布全国的产业网络，甚至新疆、西藏等边疆省份也有他们的身影。而其中，新化人与永康人的相遇（浙江永康人也曾经到全国流动修理打字机，但后来却转轨到五金生产上了），国产复印机产业的启动与失败，二手复印机之是非等，都是非常复杂和深刻的产业故事。从根本上来说，新化文印产业实际上是中国后发式现代化的缩影，其中的曲折复杂，既让人欲罢不能，又让人辗转反侧，感慨良多。这样，这篇课程论文从2006年11月到2007

年9月，我整整做了10个月，论文也从5千多字变成了5万多字。

二

在新化文印产业的研究中，我主要采用了参与观察、问卷调查和深度访谈等具体方法。

（一）参与观察

我对新化文印产业的研究开始于当时在北大校内开复印店的曾孝飞夫妇。曾孝飞是新化县洋溪镇人，1989年跟随堂叔到广东中山市学习维修机械打字机技术，由此进入文印产业。曾孝飞诚实守信，有头脑，能吃苦，一直在北大校内经营文印店到今天。在竞争残酷的北大校园内能经营20多年，这是非常不容易的。现在，他的儿子也子承父业，在长沙经营一家文印店。认识曾孝飞对于我的田野调查来说，意义重大。当时新化人在北京高校的文印产业形成了地缘网络，曾孝飞是这个网络的重要连接点，我是通过他进入了这个产业，并进入到了新化文印人的地缘网络。

当时我对曾孝飞的文印店进行了10个月的参与观察，主要围绕文印店的人员结构、经营运行和日常生活来进行。当时参与观察主要采取了集中观察和分散观察两种方式。我曾经连续多天到曾孝飞的文印店，全天观察文印店的经营运行，包括每天的业务种类、运营收入、经营中遇到的问题以及解决措施等。在基本了解了文印店的日常运行后，则采取了分散观察的方法，也就是只要有时间，就去文印店进行观察和交流。特别是在对北京地区8所高校的文印店进行普查的阶段，我更是经常和曾孝飞交流调查中遇到的问题以及如何深入进行普查等。正是通过对曾孝飞文印店的长期观察，使

得我对文印店的日常运行以及新化人的地缘网络非常熟悉，从而为后来的问卷调查和深度访谈奠定了良好的基础。由于长时间的融入，我当时已经成为了文印店的一员，由此也和曾孝飞一家结下了深厚的友谊，并一直持续到今天。

（二）问卷调查

在对新化文印产业的研究中，一个重要的问题就是新化文印人在整个北京高校内的文印市场里，到底占有多少份额？为了解决这个问题，在参与观察的基础上，我设计了北京高校文印产业调查的问卷，主要包括三个部分，一是文印店主的基本信息，包括文印店主的籍贯、年龄、性别和文化程度，重点是文印店主来自于哪个地方。二是文印店的经营运行，包括如何进入这个行业，如何学习文印技术，如何购买和维修文印设备，如何雇佣学徒等；影响文印店经营的一些具体因素，包括位置、服务态度、设备和技术水平、业务价格和多元化经营等。三是文印市场的竞争，包括校园内外的双重市场体系，价格战的发动、过程和影响等，新化文印人市场竞争的手段和方式等。当时在进行实际的问卷调查时，一般选择文印店业务较少的时间段来进行。对于新化人，主要运用地缘网络来进行；对于非新化人，则主要利用市场竞争这个话题来打开局面。当时比较深刻的感受就是，由于前期做了参与观察和访谈，对这个行业比较了解，所以问卷的问题设计相对比较内行，因此，能够引起文印店主的兴趣和共鸣，从而保证了问卷调查的顺利进行。

（三）深度访谈

经过前期的参与观察和问卷调查，我才发现新化文印产业绝不

只是一家小小的文印店，而是一个有着多年历史发展演化的庞大的产业系统。而对这个系统的深入研究，则主要是依靠深度访谈。北京大学校园内外的文印店是我研究的基点所在，先后进行了15个深度访谈。清华大学由于和北京大学处于同一个市场圈，我对清华大学内的3个文印店主也做了深度访谈。在追溯北京大学校内文印店的历史时，对于曾经在北大开文印店后搬到上海的3人进行了深度访谈，这样就获得了21个访谈个案。在对8所院校普查时，除了北京大学之外，对其他7所院校各选取校内和周边一家文印店店主进行深度访谈。同时为了弥补抽样的地域不均衡，又选取了东城区的中央音乐学院和丰台区的首都医科大学内的文印店做了深度访谈，这样就获得了16个访谈个案。对于新化文印产业的发展历史和各个阶段的关键人物，则进行了16个电话访谈。同时，还对天津复印机研究所事业发展部部长、邯郸汉光办公自动化耗材有限公司总经理、南京田中机电总经理和新化县驻北京办事处主任等4人进行了访谈。这样访谈汇总，一共获得了57个深度访谈个案。

当时的深度访谈类型主要是生命史或生活史访谈，也就是通过从业者的个人生命史来考察产业的发展史。访谈一般从父辈开始问起，然后顺着个人生命史展开，包括教育情况、就业情况、如何进入文印产业、在文印产业的发展、遇到的问题以及解决的措施、经营情况和未来展望等。访谈地点和时间以方便为宜，一般在文印店、茶馆、饭店等地方进行，大多都是一对一谈话，一个人一般要谈半天。对于从业时间长、有思想和见地的，则会连续谈好几天，直至把需要的行业信息谈完为止。当时我的角色更像是一位矿工，就是要通过深度访谈把行业宝藏都谈出来。

三

在研究的过程中，我也遇到了不少障碍和挑战，比较大的就是研究的重新设计、研究的深入以及远距离访谈等问题，后来都想办法克服和解决了，但也留下了不少遗憾。

（一）研究的重新设计

按最开始的想法，这个研究就是为了完成一篇5000字的课程作业，因此，开始的研究设计就是在北京大学校内选择3—5家文印店，进行观察和访谈，研究主题则围绕文印店的运行来展开。但在作业初稿被导师否定之后，整个研究面临着重新设计的问题。

经过与导师和同学们的交流，研究主题重新修改为新化文印现象的形成，其中的关键，就是要搞清楚新化人开的文印店在北京高校所占的市场份额，这是个定量问题，对于我这个历史专业出身，没有学过统计分析，也没有做过量化研究的研究者来说，显然是个不小的挑战，我决定向从事量化研究的同学求助。

当时我宿舍对面住的就是北京大学人口研究所搞量化研究的王存同博士，一次在和他聊天中，我提到了新化文印产业研究中遇到的市场占有率问题。他建议说，可以把北京地区的高校作为样本总量，进行等距抽样，然后对抽中的高校内的文印店进行普查，这样就可以计算出新化人在文印市场的市场占有率。后来我就按照他的建议，具体细化了研究设计，以当时北京地区的56所高校为样本总量，以7为单位进行等距抽样，共抽取了8所高校进行普查：北京大学（海淀），北京化工大学（海淀），对外经济贸易大学（朝阳），北京理工大学（海淀），首都师范大学（海淀），北京电子科

技学院（丰台），北京机械工业学院（海淀），北京第二外国语学院（朝阳）。在对这8所高校文印进行市场普查的基础上，统计出新化人在北京高校文印市场的占有率为65%。

（二）研究的深入

在研究的过程中，我逐渐发现新化文印产业经过几十年的发展，已经形成了包括服务、贸易和制造等多个产业链组成的复合产业链，而高校内的文印店只是这个产业末端的一部分。当时我收集的资料已经足够写一篇作业，所以当时面临的选择就是，研究还要不要深入下去？

经过考虑之后，我决定继续做下去，随着研究的深入，我发现新化文印产业形成了一种与产业集群相反的经济现象：产业扩散，而且是一种地缘型产业扩散。经过几十年的演化变迁，新化人发展出了遍布全国的文印产业经营网络，形成了国际贸易+专业市场+专业店的产业经营模式。

而在产业发展的过程中，遇到的一个沉重话题就是复印设备国产化的失败——1990年代中期，复印机的国产化基本全军覆没，至今国内复印机产业基本没有国产品牌，完全是日美系列之天下。当时国产化失败的一个原因就是配件国产化比率过高，一提高国产化，质量很快就下来了，产业也就不行了。国产化的推进竟然杀死了国产化，这不能不说是一个极大的产业教训和悖论。面对国内巨大的文印市场，美国的三大文印上市公司也曾经陆续进入中国市场，但最后全部失败而归。在其中，新化人以二手复印机为中心的产业链，无疑大大增加了文印市场竞争的残酷性。

在研究的深入中，我发现高校内的文印店只是末端市场的一部

分，每个高校周边还有一个边缘文印市场。同时，除了高校，新化人的文印店还大量分布于街面上，其中不乏一些高中端图文店，对于这一部分的调查，我当时是力不从心的，方法上也是欠缺的，后来就基本没有涉及，这不能不说是一个很大的遗憾。

（三）远距离访谈

随着研究的深入，新化文印产业各个阶段的代表性人物以及贸易和制造产业链的代表性人物，都需要访谈。但这些人除了在北京的之外，还分布在上海、广州、深圳、珠海等城市，一些人当时甚至在日本和美国。如何对这些人进行访谈，显然是不小的障碍和挑战。

经过和一些新化文印人的交流和探讨，我决定采用电话的方式进行远距离访谈。具体来说就是，通过新化文印人的地缘网络，获得这些人的电话或者手机号，然后进行电话访谈。当时我北大宿舍装的还是插卡电话，我就买了很多电话卡，从而保证了电话访谈的顺利进行。

电话访谈的方式经济、高效，能够迅速获得大量有效信息，并使得整个访谈进入了一个新的阶段，不仅打破了时空限制，而且使得访谈的视野和眼界都得到提高。海外二手文印设备的使用和扩散，对于新化文印产业的影响巨大。新化文印产业中最早发现二手文印设备价值的一批人中，其中一位当时就正在日本从事二手文印设备的贸易，对他的访谈，就是通过电话进行，从而获得了不少珍贵的产业发展信息。

对于远距离访谈中遇到的问题，则采用了“系统内问题要利用系统内资源解决”的思路来进行。当时访谈时，一般10个人里面有6—7个人都是愿意接受访谈的，但也有少数人，要么是不接电

话，要么是听到介绍后直接挂掉。但是，一些行业发展的关键人物是绕不过去的，怎么办？解决的办法就是让文印产业中资历深、威望高的人来相互介绍和打招呼。当时二手复印机专业店阶段的一位关键人物，挣到钱后到贵州开矿，已经不在这个行业做了。他开始拒绝了我的访谈，后来通过邹联经的介绍，打消了顾虑，才接受了访谈。因为邹联经在行业内的名声是如雷贯耳，是一位大神般的存在，在访谈中获得他的支持是至关重要的。

由于当时北京外的访谈全部通过电话进行，而且一打就是一个多小时，导致我的耳朵因为电话辐射留下了后遗症，现在只要接电话时间略微一长，耳朵就隐隐作痛，这也是做学术研究的代价。但其实这是完全可以避免的，当时电话只要按下免提键，耳朵离话筒远一些，就不会有这个问题了。

四

新化县是产生了陈天华、成仿吾、陈正湘和罗盛教等仁人志士的热土。1799年，新化洋溪船民杨海龙发明了毛板船，创造了中国乃至世界航运史上的奇迹，成为新化人水上闯荡世界的重要工具，形成了“舍下血肉喂鱼肚，拆落骨头再撑船”的毛板船精神，并在武汉再造了另外一个新化县城。历史传承，不绝如缕，160多年后，新化人以流动维修打字机起步，发展出了遍布全国的文印产业，形成了“有城镇处，就有新化文印”的“新化现象”。其中一以贯之的，是吃得苦、耐得烦、敢为天下先的湖湘文化和梅山精神。

新化文印产业是沿着血缘和地缘的网络向全国发展扩散的，我的研究也是沿着新化人的血缘和地缘网络进行的，在其中，很多新化人都给了我无私的帮助，可以说，没有热情好客而又勤奋的新化

人的帮助，我的研究是不可能完成的。从事文印产业的新化人是真正的英雄，而我只不过是个研究者和记录者。由于当时一些访谈在电话中进行，我和一些人一直到现在都还没有见过面。

做新化文印产业的研究对于我来说，意义是很重大的。当时北京大学博士毕业规定，需要发表两篇论文才能毕业。由于我2008年到河南挂职锻炼2年，根本没有时间写论文，所以，关于新化文印产业的两篇论文就成为我毕业的救命稻草。而且让人料想不及是，2015年，根据我的论文改编的文章《为什么复印店老板都是湖南人》在微信上大量传播。当时，我的母亲刚刚去世，我也正处于丧母的悲痛中，可以说，这篇文章的传播给了我很大的慰藉。后来，人民日报发表了相关评论文章，光明日报也转发了消息。应该说，论文产生了一定影响，确实让我没有想到，也算是为新化文印产业的宣传做了一点贡献。

新化文印产业是新化人传承基础上的创造，也是改革开放的产物，并且随着时代和技术的进步而不断发展。新化县梅山文化孕育的新化人，具有吃苦耐劳、霸蛮好胜、敢为天下先的性格，这些文化因素注入血缘、地缘和业缘网络，使得新化人在全国各地安身立命和发展扩张的同时，也为我国的社会文化事业的发展，做出了实实在在的贡献。特别是各个高校内外的文印店，以极其低廉的价格，为高校的教学和科研以及学生的学习，提供了极大的便利，这是新化人为我国高等教育的发展做出的不可磨灭的贡献。

新化文印产业不断发展壮大是多种因素的结果。举其要者，一是改革开放以来国家和社会对文印产业不断增长的需求，使得这个市场越来越大。二是各级政府的扶持。随着产业的兴起，各级政府在政策、资金、人才等方面不断出台措施，全力扶持这一产业的发

展壮大。三是新化文印人的不断创新。创新是产业发展的不竭动力，新化文印产业之所以经历了多个发展阶段，新化文印人之所以不断掌握核心技术，最根本的原因就是与时俱进，追求卓越，不断创新。

新化文印产业的发展和壮大是整个国家产业发展的缩影，也是国家和社会进步的重要表现。从国内外的产业发展经验来看，新化文印产业的生命在于专业化，在于技术的持续创新和发展。新化文印产业需要借鉴“沙县小吃”等地缘性产业发展的经验，在全国市场统一规范“新化文印”品牌，形成品牌效应。同时应该加大研发投入，在文印产业的核心技术方面不断实现突破。还应该借势现代的金融和资本市场，引进风险投资，积极谋划企业的整合和上市。

正如中国革命是缝隙中的革命，最后汇成了浩浩荡荡的革命洪流一样。新化文印开始也是缝隙中的产业，顽强生长，历经磨难，经过60多年的发展壮大，最终形成了今天30多万从业人、15万家门店、产值超1000亿元、占据全国70%以上市场份额的辉煌伟业。

参考文献

艾尔·巴比，2018，《社会研究方法》，邱泽奇译，北京：华夏出版社。

陈向明，2000，《质的研究方法与社会科学研究》，北京：教育科学出版社。

L.乔金森、丹尼，2009，《参与观察法》，龙筱红、张小山译，重庆：重庆大学出版社。

潘绥铭、黄盈盈、王东，2011，《论方法：社会学调查的本土实践和升华》，北京：中国人民大学出版社。

苛费尔、斯丹纳、布林克曼、斯文，2012，《质性研究访谈》，范丽恒译，北京：世界图书北京出版公司。

袁方主编、王汉生副主编，2014，《社会研究方法教程》，北京：北京大学出版社。

高勇．有关概念测量的三个故事与一个教训[M/OL]//赵联飞，赵锋．社会研究方法评论：第2卷．重庆：重庆大学出版社．

有关概念测量的三个故事与一个教训

高勇[①]

摘要：在定量研究中，“测量”是一个看似简单实则复杂的环节。研究过程中某些草率的决策，往往导致研究者对于经验事实产生错误的判断。本文简要地介绍了三个研究实例，分别涉及到“社会封闭”“地区收入不平等”“人类发展指数”等概念的测量。在这些实例中，测量结果的真实意蕴往往偏离了研究者原本的测量意图，出现了某种“非预期后果”。本文认为，测量本身就是一种具体的社会过程，不能只从抽象概念及概念维度的操作化来理解。只有把测量还原为具体的社会过程时，我们才能透彻地理解测量。

关键词：测量；社会过程；概念；操作化；社会闭合

Abstract: In quantitative research, "measurement" is a seemingly simple but complex process. Some sloppy decisions in the research process often lead to misleading conclusions about empirical facts. This paper briefly introduces three research examples, which involve the mea-

---

①作者简介：高勇，中国政法大学社会学院教授、博士生导师。中国社会学会社会分层与流动专业委员会理事、北京社会学会理事。联系方式：gaoyongvincent@163.com。

surement of concepts such as "social closure", "regional income inequality" and "human development index". In these examples, the true meaning of the measurement results deviates from the researcher's original measurement intentions, with some "unintended consequence". This article argues that measurement itself is a concrete social process that cannot be understood only from the perspective of abstract concepts and operationalization. It is only when measurements are considered as concrete social processes that we can understand measurements thoroughly.

Key words: Measurement; Social Process; Concept; Operationalization; Social Closure

（社会科学的实际工作）不是"实验室"，而是"工作室"，是手工匠人或15世纪文艺复兴时期的画家们的那些工作室，也就是说它有着各种各样的错误，起步不当、摇摆不定、陷入困境、推倒重来，诸如此类，不一而足。研究者们尽管各自工作的进展程度不一，但都会把自己力图建构的对象摆出来，接受其他所有人的质疑与批评。这些人的行动方式，用传统行话来说，就像这行当里的老同事、老伙计，他们会各尽所能，把他们通过过去所有的尝试和失败教训而逐渐积累起来的集体性经验贡献出来。

——布尔迪厄，《反思社会学导引》

在几乎所有的社会研究方法教材中，测量都是非常重要的章节。它是"假设—演绎"研究范式中的重要一环，是连接理论层面与经验层面的关键节点。在定量研究中，"测量"是一个看似简单实则复杂的环节。研究过程中某些草率的决策，往往导致研

究者对于经验事实产生错误的判断。本文试图通过几个研究实例的讨论，说明测量环节可能存在的研究陷阱，指出认真考察测量过程的必要性。

一、“社会闭合”

（一）概念的提出

科尔曼（Coleman，1988）在《人力资本生成中的社会资本》一文中提出了他的“社会资本”概念：它就是某种社会结构，有着便利一些行动者采取某种行动的功能后果。[①]缺了这种特定的社会结构，那些行动者的目标就没法达成。例如，某个行业是由内部有亲缘或地缘关系的小圈子构成的，这种特定的社会结构便利了他们做这一行业的生意，这就构成了一种社会资本；某个城市中有着邻里之间彼此照看的规范，这种特定的社会结构便利了他们可以让小孩自由在街上玩耍而不必费心监管，这也就构成了一种社会资本。

与以往的“非正式制度”“正式制度”“信息渠道”等概念相比，“社会资本”增加了哪些理论要素？第一，所有的“非正式制度”“正式制度”“信息渠道”都需要落实到某种实在的社会结构上，如有彼此互惠义务的群体、有共享规范和规则的组织、互通信息的网络圈子等等。第二，如同资本具有一定程度的“通用性”一

①不过可以想见，这种社会结构同时也一定阻碍了另一些行动者采取另一些行动。在这种意义上，特定社会结构对一些人采取一些社会行动是“资本”，对另一些人采取另一些社会行动来说就是“妨碍”。有利于父母管教孩子的一种社会结构，对父母是“社会资本”，但对那些顽皮孩子来说，无疑是一种“社会妨碍”。科尔曼本人也明确地意识到这一点，但并未展开讨论（Coleman，1988：S105）。

样，原本用于某种目标而形成的社会结构，随后完全可以成为服务于另一目标的手段，它同样具有“通用性”。正是在这个意义上，那些社会结构才具有“资本”的特性。

人际关系网络就是这样一种结构。不过，它要有“通用性”，就必须具有“社会闭合”（social closure）的特点。你与工作同事有关系网络，你的孩子与同学有关系网络，但是如果你的孩子的同学正好是你工作同事的孩子，那么关系网络就“闭合”了。此时，原来因工作建立起来的关系网络就可能被挪用来监督和引导孩子，它便利了你的这一目标。这种人际关系网络有了通用性，因此它构成了父母的一种社会资本。在这样的社会结构中，人们之间的联结不是单一性的，而是多重联结（a multiplex relation）；在某一背景下形成的联结可以非常方便地挪用到另一背景当中。现代社会中，这种多重联结的人际关系网络对于社会的运作发挥着至关重要的作用。

与“社会资本”概念相比，附属概念“社会闭合”有具体的指向，功能论的色彩也少一些。它强调了人际关系中多重联结的结构性作用，为社会结构研究增加了新的维度。某些研究者甚至认为，“社会闭合”这一概念更具有原创性（Carbonaro，1998）。

（二）经验测量与证据

科尔曼的论文尽管有数据佐证，但是仍然重在提出思想，而不是经验测量。他以往一直在从事教育领域的研究，因此就顺带把这一思想放在教育领域中来阐释。由于手头的数据中并没有对于“社会闭合”的直接测量，科尔曼使用了“由于搬家而导致的转校次数”来作为对这一概念的近似指标。这一做法不能令人信服，有严

重的“贴标签”嫌疑。另一个证据是宗教性私立学校的辍学率远低于公立学校。宗教性私立学校中，学生父母的人际关系网络可能更具有多重联结的特点：既是教友，又是好友的父母。但这一证据也太于单薄，因为我们可以想象出很多备择假设来解释同一事实。

有可能受限于科尔曼当初的例证，有关“社会闭合”的后续研究基本都是在教育领域进行的，虽然这一概念本身有更深远的意义。在1988年美国教育纵贯研究（NELS：88）的第二轮追踪调查中[①]，研究者特意对这一概念进行了操作化，设计了如下题项来测量“代际关系闭合”：

53. 你能说出你孩子的一个好朋友的名字（或外号）吗？
能 ……………1（继续回答54题）
不能 …………2（跳答56题）

54. 请列出你孩子的好朋友的名字（或外号），然后说明：
（A）他（她）和你孩子是否在同一所学校上学？
（B）你是否认识他（她）的家长？

好朋友的名字	A 同一所学校上学			B 认识家长	
	是	否	不适用	是	否
1.	1……	2……	3	1……	2
2.	1……	2……	3	1……	2
3.	1……	2……	3	1……	2
4.	1……	2……	3	1……	2
5.	1……	2……	3	1……	2

①NELS的基线调查是1988年实施的，第二轮追踪调查是在1992年实施的。读者现在仍然可以从以下网址下载到当年的问卷：https：//nces.ed.gov/surveys/nels88/questionnaires.asp。

上述54题B中回答“是”的数目，就可以作为“社会闭合”的一种测量指标。它代表了家长认识多少位别的家长，这和科尔曼当初举的例子是吻合的。从概念操作化的角度来看，这似乎没有任何问题。基于NELS数据，研究者对于科尔曼的理论假说进行了经验检验（Carbonaro，1998；Morgan & Sørensen，1999）。卡博纳罗发现，在个体层面社会闭合对于学生的数学成绩的提升确实有正面影响。摩尔根和索伦森的研究使用多水平的随机效应模型，发现在学校层面上“社会闭合”对学习成绩基本上没有什么正向作用。

（三）我们究竟测量的是什么？

然而，这里面有一个最基本的问题需要澄清：我们测量出来的，究竟是什么？霍里南和库比切克（Hallinan & Kubitchek，1999）指出，这一变量很可能与“社会闭合”关系不大。

让我们回到事实层面，来仔细考察一下测量过程是怎么发生的：首先，被访者要列出他们孩子的朋友名字；然后，我们才继续提问有关这些朋友的进一步信息。第一步给后面的步骤定了上限：如果你只能列出孩子的3个朋友，那么就算你的社会关系网络都是闭合的，那你在后面的步骤中也最多只能说自己认识孩子的3个朋友的父母；如果你能列出孩子的5个朋友，那么就算你的社会关系网络并不太闭合，你在后面的步骤中也有很大概率说出自己认识孩子的3个朋友的父母。

变量的大部分变异，究竟是来自“知道的小孩朋友数”，还是来自“社会关系闭合度”？数据表明，前者更符合事实。列出小孩名字的个数，与列出父母名字的个数是强相关，相关系数接近0.75。家长平均能够说出孩子的4个好朋友的名字，平均能够说出

孩子的3.3个好朋友的父母名字。基本上来说，知道自家小孩的朋友名字，一般就能够认识他们的父母。

于是，我们了解到的信息，与其说是“社会关系闭合度”，不如说是“孩子的人缘”（Hallinan & Kubitchek，1999）。我们自以为看到的是，社会关系闭合度高的小孩成绩更好。但事实可能是，成绩好的孩子人缘就比较好，因此父母知道的朋友数也越多；这就同社会闭合度没有关系。或者，我们自以为看到的是，社会关系闭合度与成绩没有关系；事实上我们证明的只是，孩子的人缘和成绩没有关系。

还有另一种可能：我们了解到的信息，与其说是“社会关系闭合度”，不如说是拉鲁（2018）在《不平等的童年》中指出的“协作培养方式”（concerted cultivation）。父母越是与孩子多交流，他就越了解孩子的生活细节，能说出更多的孩子的朋友名字。当然，这种培养方式下，孩子的学业成绩也往往更高。无论如何，这和社会闭合度也没有什么关系。[①]

在这里，测量出现了问题。然而，如果简单沿用教科书式的“假设—演绎”模型，我们并不太容易发现这里的问题。概念与测量之间的脱节，比我们以为的更容易发生（Martin，2017：38-40）。

①在经历了这些学术争论之后，2002年和2004年美国的ELS（Education Longitudinal Study）调查重新设计了社会网络模块的提问方式。利用新的数据，摩尔根和陶德（Morgan & Todd，2009）再次对“社会闭合”效应进行了检验。结果发现，在控制了家庭背景、学生交往网络等因素后，社会闭合在公立学校中对成绩并没有明显促进作用，但在天主教学校中对成绩有一定促进作用。

二、"收入不平等"与幸福感

(一)理论命题的提出

发展经济学家赫希曼(Hirschman & Rothschild, 1973)在《经济发展过程中对收入不平等容忍程度的变化》(*The Changing Tolerance for Income Inequality in the Course of Economic Development*)一文中提出了著名的"隧道效应"。人们对于不平等的容忍程度是随着发展的不同阶段而变化的。在发展的早期阶段，收入差距逐渐拉开，但是人们对于不平等具有相当的容忍程度。因为人们看到身边的高收入者时，会感觉到这提供了生活改善的希望与前景。如同处在交通堵塞的隧道中的司机，他在看到旁边车道开始移动时，即使自己还没有移动也会充满希望一样。收入差距的适度增大，增强了人们对于未来生活福祉提升的预期，从而提升了他/她的幸福感。但是在早期阶段过后，人们对不平等的理解方式就改变了。如同处在交通堵塞隧道中的司机，在看到旁边车道一直移动，自己却迟迟不能起步时一样会充满抱怨。

赫希曼的这一理论命题有着非常重要的社会学意味。他表明，人们对收入不平等的容忍程度，并不是一种稳定的个体偏好，而是受到对外在社会情境的"定义"的影响。我们可以把它看成是社会学中"情境定义"命题在宏观经济发展中的应用。"情境定义"并不只是一个在微观人际互动中适用的命题，它完全可以推广到我们对宏观社会变迁的理解中。

(二)经验测量与证据

在中国高速发展的背景下，有很多研究者使用调查数据来讨论

收入不平等程度对于人们心态观念的影响。利用全国综合抽样调查数据（如CGSS）计算得出区县（甚至更小单位的）基尼系数（或泰尔指数等相似指标），然后引入回归模型进行分析，这是一种讨论收入不平等对其他因素影响的常见做法。研究者并没有得到一致的答案：有的发现地区基尼系数对幸福感有正向作用，有的发现有负向作用，还有的发现两者存在一种"倒U性曲线关系"。

我们用CGSS2013数据来看一下区县基尼系数与各区县幸福感均值的关系。初步的证据（图1）表明，两者之间似乎存在着一种正向关系：区县基尼系数越大，幸福感的评价值越高。这似乎也可以印证赫希曼的理论命题。为了直观起见，我们在此只用散点图来展示双变量关系，但是在更复杂的统计模型分析中，收入基尼系数对幸福感的效应也非常显著。

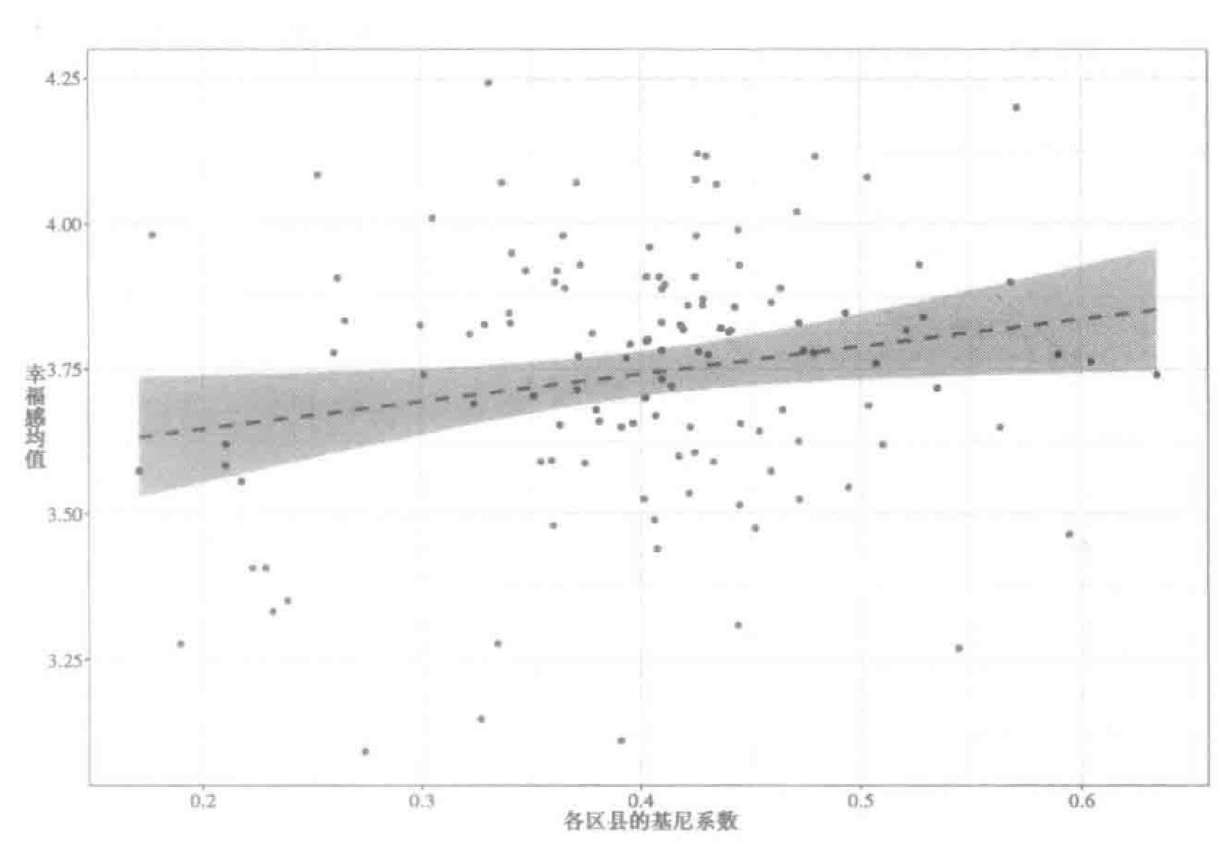


图1　区县基尼系数与幸福感均值

（三）我们究竟测量的是什么？

但是，我们对于基尼系数的测量似乎有些不太对头。让我们把图1中最左端和最右端那些散点所代表的区县所在省份标注出来，由此得到图2。为什么那些基尼系数最低的区县都来自于北京、天津、广东？为什么那些基尼系数最高的区县却来自于云南、福建、安徽、山东？这既和我们的日常经验不相符合，也和赫希曼的理论推断不相符合。

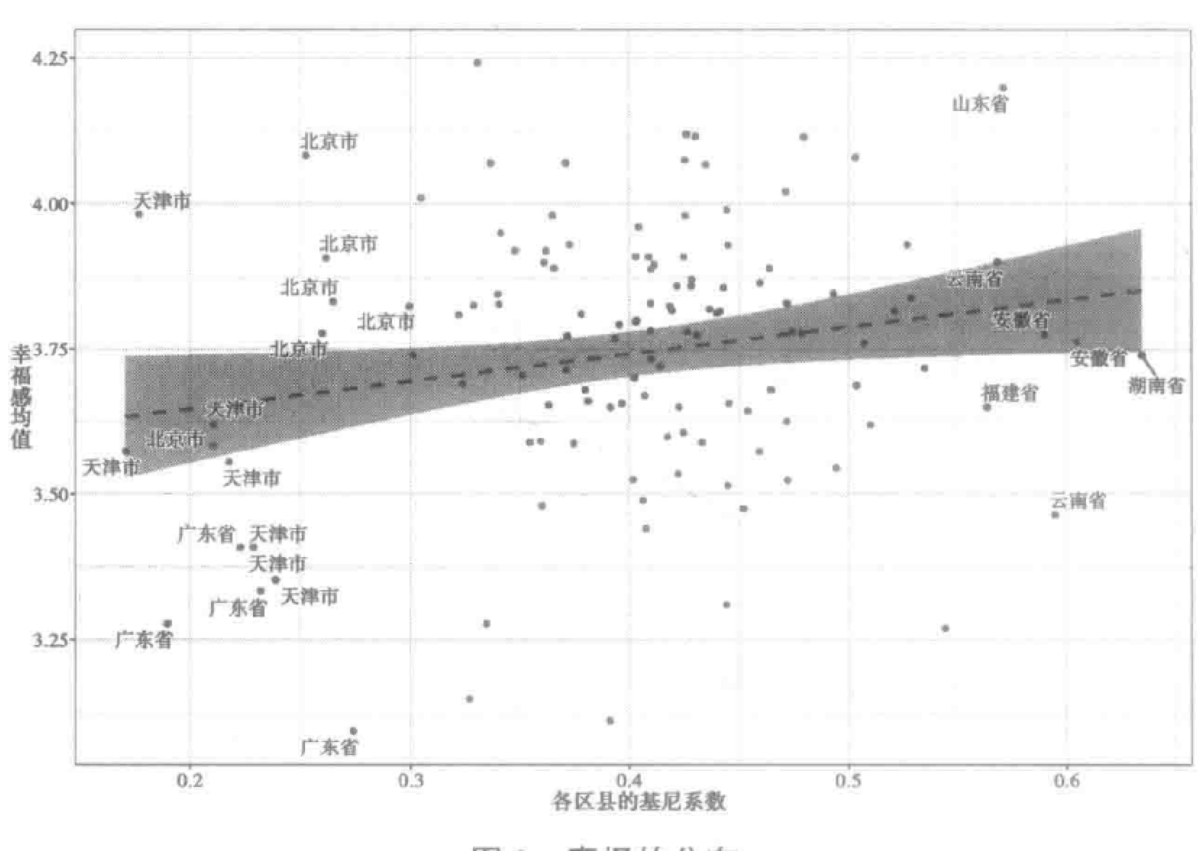


图2　奇怪的分布

原因出在了我们对基尼系数的测量方式上，而这又和CGSS调查的抽样方式有关。CGSS调查的抽样有必选层与抽选层的设计。[①]

①更多细节请参阅《中国综合社会调查（CGSS）第二期（2010-2019）抽样方案》（中国综合社会调查网，http：//cgss.ruc.edu.cn/xmwd/cysj.htm）。

北京、上海、天津、广州、深圳五个大城市是必选层，必选层的初级抽样单元是街道：首先抽取40个街道，每个街道中抽取2个居委会，每个居委会抽取25个家庭户。因此，大城市的样本全部来自居委会，多数区县仅抽到2个居委会，这2个居委会一定在同一个街道。此外，由于城市房产价格的原因，大都市地区居住于同一居委会中的居民家庭收入往往较为接近。基于上述原因，CGSS必选层区县内的样本具有很高的同质性，因此测算出来的收入基尼系数一定较小。

其余全国所有家庭户构成抽选层，抽选层的初级抽样单元是区县：首先抽取100个区县，每个区县中抽取4个居委会或村委会，每个居委会或村委会抽取25个家庭户。在这样区县样本中，既有居委会也有村委会（根据城市化水平进行配比），所有区县都被抽中了4个村委会或居委会，这4个村委会或居委会未必在同一街道或乡镇。此外，在农村地区，隶属于同一村委会的村民因收入提升而搬离本村庄的可能性固然存在，但相对城市而言较小。基于上述原因，CGSS抽选层区县内样本异质性程度更高，因此测算出来的收入基尼系数一定更大。

这样一种抽样方式导致了一个奇特的“非预期后果”：测算出来的区县收入基尼系数与区县城市化水平有很强的相关性。城市化水平越高，测算出来的区县收入基尼系数越低。在这种情况下，我们测量的与其说是区县的“收入不平等程度”，不如说是区县的“城市化水平”。我们以为看到的是，收入不平等程度越高，主观幸福感越高；而我们真正看到的有可能只是，大城市中的居民的主观幸福感较低。事实上，在区分了必选层和抽选层之后，回归模型中

收入基尼系数对幸福感的效应就大大衰减，失去了显著性。[1]

以“基尼系数”的公式测算出来的数字，并不一定测量的就是“收入不平等程度”。数据的抽样方式以隐蔽的方式扭曲了我们的测量。

三、“人类发展指数”

（一）新的发展理念

早期对于发展的看法，将它等同于经济增长，同于财富增长。阿玛蒂亚·森（2002）在《以自由看待发展》中提出了不同的理念：“发展可以视为一种扩张人们享有的真实自由的过程。它的焦点是人类的自由，并与发展的狭隘观点形成强烈对比。这些狭隘的观点，像是把国民生产总值的成长、个人所得的增加、工业化、技术进步或社会现代化，看成是发展。国民生产总值或个人所得的增加，对于扩张社会成员的自由而言，当然可以是很重要的工具。但是自由也取决于其他因素，譬如社会和经济的安排以及政治和公民的权利。”这种发展观也被称为发展的“可行能力”视角。

联合国开发计划署（UNDP）编写的《1990年人类发展报告》中将这种理念明确化，正式提出了“人类发展”的概念，以与以前的发展观相区分：“人是国家真正的财富。发展的基本目标是创造有利环境，使人类长寿、健康和富有创造力……人类发展是一个不断扩大人们选择权的过程。最关键的选择权包括拥有健康长寿的生命、受教育和享受高品质生活的权利。其他选择权包括政治自由、

①有兴趣的读者可以在如下网址 http://gao-yong.cn/查看更详尽的细节。

有保障的人权与自尊……人类发展将商品生产与商品分配、拓展人类能力与利用人类能力结合起来。人类发展也关注选择权——关注人们为了确保生计应该有什么，是什么，做什么。此外，人类发展不仅与满足人类基本需要有关，同时也是一个动态参与过程。无论是在欠发达国家还是在发达国家，都是如此。”

（二）新的测量方式

在提倡新理念的同时，联合国开发计划署（UNDP）也提出了测量发展的新指标：人类发展指数。人类发展指数是三个分项指数的综合：健康指数（人均预期寿命）、教育指数（文盲率与平均受教育年限）、人均收入指数（以平价购买力计算的人均收入）。它对应于发展的三种基础性能力：长寿而健康的生命、获得知识的渠道、有尊严的生活水准。三个分项指数进行标准化之后的几何平均数，就是最终的人类发展指数。其中的技术细节，可以参看UNDP的说明。人类发展指数的推出，对于改变国际社会中的发展理念，更准确地测度某一国家的发展水平，起到了积极的作用。

这一指数随后也被用来测量某一国家内部不同地区的发展水平。1997年开始，UNDP发布的《中国人类发展报告》也开始用这种方式来测量中国各省、直辖市和自治区的“人类发展指数”。这一指数被认为超越了单纯用GDP来衡量发展的旧观念，能够促使各省更加重视发展的非经济维度。图3显示了2017年中国各省的人类发展指数情况（联合国开发计划署等，2019）。

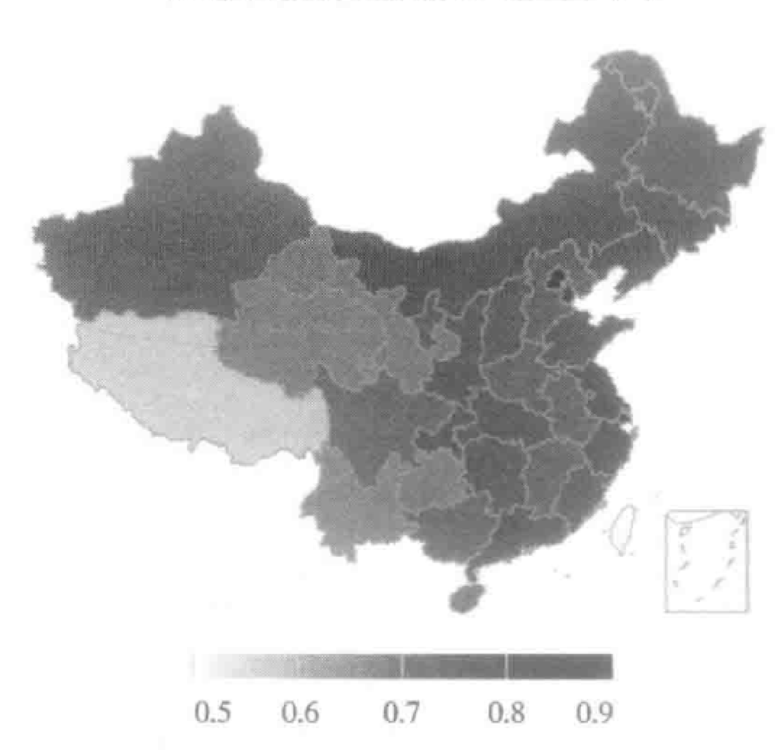


图 3　中国各省人类发展指数（2017年）

资料来源：联合国开发计划署、清华大学中国发展规划研究院、国家信息中心联合编制，2019，《中国人类发展报告特别版——历史转型中的中国人类发展40年：迈向可持续未来》，中国出版集团、中译出版社。

（三）我们究竟测量的是什么？

在现实当中，人们对于各个省人均发展指数的具体数值并不敏感，各省真正在意的是它们在人均发展指数上的排名。我们把各省在人类发展指数上的排名与他们在人均收入上的排名绘图，结果呈现在图4中。图4显示，尽管人类发展指数的构建中包含了除人均收入之后的其他两个指标（健康与教育），但是它的排名与收入排名是强相关。原因在于，各省在预期寿命和受教育年数上的差异并没有收入那样大；预期寿命和受教育年数本身与人均收入也是强相关。

尽管在理论上人类发展指数这一指标确实是诸多维度的综合加总，但是在这一具体情境中，它对案例的区分与按人均收入进行的区分基本类似。如格尔曼等所言，在这一意义上我们测量的其实主要还是“人均收入”，虽然它经过了一些“神秘的转换”而获得了更加“炫目的名称”（Gelman，Hill & Vehtari，2021：22）。

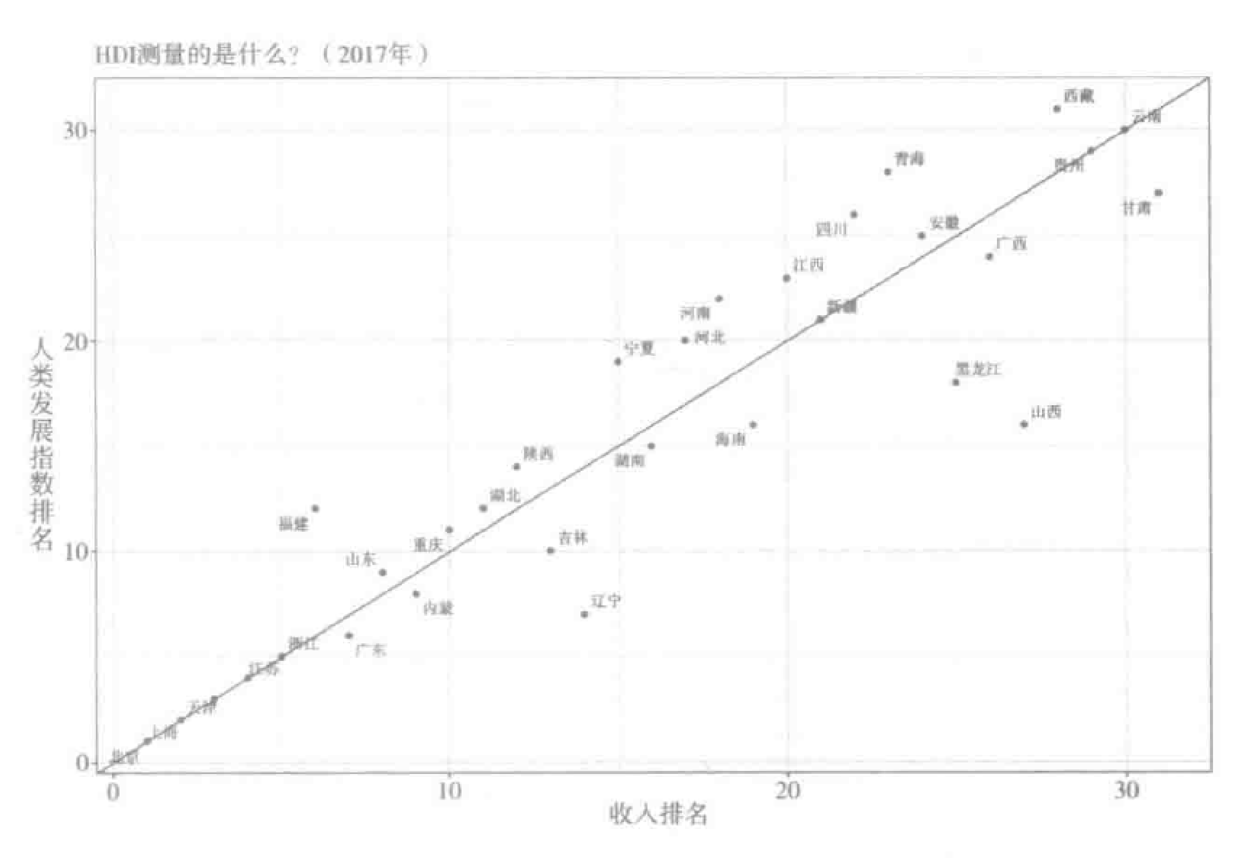


图 4　2017 年各省人类发展指数的排名与收入排名

四、一个教训

社会科学方法首先要克服的问题，就是各种形式的“自欺”（赵锋，2019）。在把概念与经验观察相关联的过程中，我们尤其要防范各种“自欺”。研究本身就是一个社会过程。我们要想了解作为研究对象的“大”的社会过程，前提是首先要充分理解自身研究

这一“小”的社会过程。社会学要成为科学，必须要用社会学的眼光来看待自己的社会研究过程。只有真正理解研究过程中究竟发生了什么，我们才能够免于各种“自欺”。

测量的通常流程是从概念出发，逐步进行概念化与操作化，最后得到反映概念要素的变量指标。“假设—演绎”的研究过程强调对于理论的检验，因此研究者容易急于去证明自己心仪的某些理论是正确的，而容易忽略测量背后潜藏的一些基本事实。由于多种原因，我们意图测量的与实质测量的内容未必是一致的。测量出来的信息究竟是什么，不仅是理论问题，也是经验问题。验证理论很重要，但前提是先要搞清楚基本事实。有必要回到事实层面，原原本本地去看数据，去琢磨现实生活中的社会过程可能是怎么回事，而不是只停留在抽象概念上。

为了防止测量过程中的种种失误，我们有必要在测量完成之后，去认真考察这一变量与其他各种变量的相关结构。如果这一变量与其他某个理论上完全独立的变量呈现出较强相关关系时，我们就需要警惕测量是否出现了某种偏差。此时，无论我们怎样为理论进行辩护，如果测量结果可以被解读为另外一种意义，其结论也就不再牢靠了。

参考文献

布尔迪厄、华康德，2015，《反思社会学导引》，李猛、李康译，北京：商务印书馆。

拉鲁，2018，《不平等的童年：阶层、种族和家庭生活（第2版）》，宋爽、张旭译，北京：北京大学出版社。

联合国开发计划署、清华大学中国发展规划研究院、国家信息中心联合编制，2019，《中国人类发展报告特别版——历史转型中的中国人类发展40年：迈向可持续未来》，北京：中国出版集团、中译出版社。

森·阿玛蒂亚，2002，《以自由看待发展》，任赜、于真译，北京：中国人民大学出版社。

赵锋，2019，《科学作求知激情，方法为研究技能》，《中国社会科学报》3月27日。

Abbott, Andrew 2004, *Methods of Discovery: Heuristics for the Social Sciences*. New York: Norton.

Carbonaro, William J. 1998, "A Little Help from My Friend's Parents: Intergenerational Closure and Educational Outcomes." *Sociology of Education* 71(4):295-313.

Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94: S95-120.

Gelman, Andrew, Jennifer Hill & Aki Vehtari 2021, *Regression and Other Stories*. New York: Cambridge University Press.

Hallinan, Maureen T. & Warren N. Kubitschek 1999, "Conceptualizing and Measuring School Social Networks." *American Sociological Review* 64(5):687-93.

Hirschman, Albert O. & Michael Rothschild. 1973, "The Changing Tolerance for Income Inequality in the Course of Economic Development." *Quarterly Journal of Economics* 87(4):544—66.

Martin, John Levi 2017, *Thinking Through Statistics*. Chicago: University of Chicago Press.

Morgan, Stephen L. & Aage B. Sørensen 1999, "Parental Networks, Social Closure, and Mathematics Learning: A Test of Coleman's Social Capital Explanation of School Effects." *American Sociological Review* 64(5):661-81.

Morgan, Stephen L. & Jennifer J. Todd 2009, "Intergenerational Closure and Academic Achievement in High School: A New Evaluation of Coleman's Conjecture", *Sociology of Education* 82(3): 267-286.

《社会研究方法评论》征稿启事

《社会研究方法评论》是以集刊方式创办的专业学术交流出版物，由中国社会科学院社会学研究所社会调查与方法研究室社会调查与数据处理研究中心负责编辑，每年2~4期，采用电子书形式，并委托重庆大学出版社、重庆大学出版社出版。目前，刊物开设“论文”“方法前言”“调查报告”“田野与课堂”“铸器坊”“新知与书评”等栏目，热诚欢迎国内外学者投稿。投稿注意事项如下：

1.《社会研究方法评论》（srmr@cass.org.cn）为本刊唯一投稿渠道，编辑部不接受邮寄纸质投稿。

2.本刊刊登以社会研究方法为主题的各类学术论文和研究述评等，不刊登国内外已公开发表的文章（含电子网络版）。学术论文以15000~20000字为宜，其余稿件以3000~30000字为宜。

3.稿件第一页应包括以下信息：（1）文章标题；（2）作者姓名、身份证号码、单位、联系电话、通信地址、E-mail等。

4. 稿件第二页应包括以下信息：（1）文章标题；（2）中文提要（不超过200字）；（3）3~5个中文关键词；（4）文章的英文标题、作者姓名的汉语拼音（或英文）；（5）英文提要（不超过200个单词）。

5.投稿文责一律自负，凡采用他人成说务必加注说明。在引文后加括号注明作者、出版年份及页码，详细文献出处作为参考文献列于文后，以作者、出版年份、书名（或文章名）、译者、出版地点、出版单位（或期刊名）排序。文献按作者姓氏的第一个字母依A-Z顺序分中、外文两部分排列，中文文献在前，外文文献在后，

外文文献中的书名及期刊名用斜体，论文题目写入“ ”内。作者本人的说明性文字均采用当页脚注。

6. 文章正文的文内标题、表格、图、公式以及脚注应分别连续编号。一级标题用编号一、二、三……二级标题用（一）、（二）、（三）……三级标题用 1.、2.、3.……四级标题用（1）、（2）、（3）一级标题居中，二级标题左对齐，三级标题左缩进两格。

7. 每张表格、图均应达到出版印制要求，并置于文中恰当位置。

8. 本刊取舍稿件唯以学术为标尺，实行匿名评审稿件制度，未来将实行三审定稿制。

9. 文章一经发表，版权及著作使用权即归《社会研究方法评论》编辑部。凡涉及国内外版权及著作使用权问题，均遵照《中华人民共和国著作权法》及有关国际法规执行。本刊刊登的所有文章如需转载或翻译，须经《社会研究方法评论》编辑部书面授权。

10. 本刊刊登的所有文章的网络版版权及著作使用权归《社会研究方法评论》编辑部所有。

11. 请勿一稿数投。投稿 3 个月未收到刊用通知者请自行处理。来稿一经采用，即奉稿酬和当期刊物电子版。

12. 本刊热诚欢迎国内外学者将已刊行或出版的学术论文或专著赠予本刊编辑部，亦热诚欢迎国内外学者或机构将社会方法研究领域的重要学术信息及时通报给我们，以期将《社会方法研究评论》建设成学术交流的平台。

13. 本刊不收取任何版面费，如发现违规行为可拨打举报电话：010-85195567。